REPUBLIC OF TÜRKİYE
ÇANAKKALE ONSEKİZ MART UNIVERSITY
SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF FOREIGN LANGUAGES EDUCATION
ENGLISH LANGUAGE TEACHING PROGRAM

EXPLORING ENGLISH TEACHERS' AND STUDENTS' EXPERIENCES ON THE IMPLEMENTATION OF MONE'S NEW ASSESSMENT AND EVALUATION REGULATION TOWARDS ENGLISH LANGUAGE COURSES

MASTER'S THESIS

SAİD YASİR KARAMAN

Supervisor
ASST. PROF. SEDAT BECEREN

ÇANAKKALE – 2025

REPUBLIC OF TÜRKİYE
ÇANAKKALE ONSEKİZ MART UNIVERSITY
SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF FOREIGN LANGUAGES EDUCATION
ENGLISH LANGUAGE TEACHING PROGRAM

EXPLORING ENGLISH TEACHERS' AND STUDENTS' EXPERIENCES ON THE IMPLEMENTATION OF MONE'S NEW ASSESSMENT AND EVALUATION REGULATION TOWARDS ENGLISH LANGUAGE COURSES

MASTER'S THESIS

SAİD YASİR KARAMAN

Supervisor
ASST. PROF. SEDAT BECEREN

ÇANAKKALE – 2025

T.C.
ÇANAKKALE ONSEKİZ MART ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

Said Yasir KARAMAN tarafından Dr. Öğr. Üyesi Sedat BECEREN yönetiminde hazırlanan ve **24/11/2025** tarihinde aşağıdaki jüri karşısında sunulan "**Exploring English Teachers' and Students' Experiences on the Implementation of Mone's New Assessment and Evaluation Regulation Towards English Language Courses**" başlıklı çalışma, Çanakkale Onsekiz Mart Üniversitesi Lisansüstü Eğitim Enstitüsü **Yabancı Diller Eğitimi Anabilim Dalı**'nda **YÜKSEK LİSANS TEZİ** olarak oy birliği ile kabul edilmiştir.

Jüri Üyeleri	**İmza**
Dr. Öğr. Üyesi Sedat BECEREN (Danışman)	……
Prof. Dr. Salim RAZI	……
Doç. Dr. Orhan KOCAMAN	…….

Tez No : 10738348
Tez Savunma Tarihi : 24/11/2025

……
Prof. Dr. Melis ULU DOĞRU
Enstitü Müdürü
.../.../20…

ETHICAL STATEMENT

This thesis has been prepared in accordance with the Thesis Writing Guidelines of the Graduate Education Institute of Çanakkale Onsekiz Mart University. I confirm that all the data, information, and documents presented in this study were obtained in line with academic and ethical principles. I have presented all findings, evaluations, and results in compliance with scientific ethics and academic integrity. All sources that I benefited from have been properly cited, and no changes have been made to the original data. I also declare that the study presented in this thesis is original. In case of any contrary situation, I accept full responsibility for any possible consequences that may arise.

ETİK BEYAN

Çanakkale Onsekiz Mart Üniversitesi Lisansüstü Eğitim Enstitüsü Tez Yazım Kuralları'na uygun olarak hazırladığım bu tez çalışmasında; tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi, tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu, tez çalışmasında yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi, kullanılan verilerde herhangi bir değişiklik yapmadığımı, bu tezde sunduğum çalışmanın özgün olduğunu, bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi taahhüt ve beyan ederim.

Said Yasir KARAMAN
24/11/2025

ACKNOWLEDGEMENTS

This study was supported by the Scientific and Technological Research Council of Türkiye (TÜBİTAK) 2210-A National Graduate Scholarship, which significantly supported my academic journey.

First and foremost, I would like to express my sincere gratitude to everyone who supported me throughout the process of completing this thesis.

I am deeply grateful to my supervisor Asst. Prof. Sedat BECEREN for his invaluable guidance, continuous support, and insightful feedback at every stage of this thesis. His expertise, encouragement and patience were essential for the completion of this study.

I would also like to thank the members of the thesis committee, Prof. Dr. Salim RAZI, Doç. Dr. Orhan KOCAMAN for their constructive comments and valuable contributions to this research.

My sincere thanks to the Department of English Language Teaching at Çanakkale Onsekiz Mart University for providing a supportive academic environment and education.

My sincere thanks also go to the English teachers and seventh-grade students who generously participated in this study. Their insights and assistance and openness made this research possible.

I am forever thankful to my family Aysel KARAMAN, Hatice ATEŞ, and Zekayi KARAMAN for supporting me in every way possible at every step of my study. I dedicate my studies to them.

I would like to thank all my classmates in the ELT Master's Program for their sincere friendship, shared experiences, and encouragement throughout this challenging yet rewarding journey. Studying and growing together with such a supportive group was a privilege that made the process more meaningful. In particular, I am grateful to my friends Emre UYGUN, Gökçe Su BALKAN GÖĞÜSGEREN, Deniz GÜL for their insights helped me see things more clearly. My special thanks to my dear friend, Özge Can CEYLAN, who stood by me during the most stressful and uncertain moments, with unwavering support and empathy. I am truly fortunate to have had such a friend throughout this journey.

I am also deeply grateful to the wonderful people at Walpurga Cafe, which became more than just a place to work, it felt like a second home. The warm-hearted family behind it welcomed me with kindness and treated me as one of their own. Their support, quiet encouragement, and comforting presence during the long hours of writing held me up more than they probably know. This thesis carries the memories of every cup of coffee, every word written, and every moment shared in that space.

I would like to extend my heartfelt thanks to Kayla BLIGHT, who, despite being miles away, remained one of the closest people in my life throughout this journey. Her genuine friendship, attentive ear, and constant encouragement carried me through both the most stressful and the most joyful moments of my academic life. She has witnessed every stage of this academic journey not just as a friend, but as a constant source of strength, warmth, and perspective. Thank you, Kayla, for being the kind of friend who makes distance irrelevant and support unforgettable.

This thesis may carry my name, but it was shaped by the presence, support, and kindness of many people along the way. Every moment of doubt was eased by a voice of encouragement, and every step forward was made lighter through shared strength. For all those who stood beside me, near or far, silently or actively, I offer my deepest gratitude. This journey was mine, but it was never walked alone. And for that, I am truly thankful.

Said Yasir KARAMAN

Çanakkale, November 2025

ABSTRACT

EXPLORING ENGLISH TEACHERS' AND STUDENTS' EXPERİENCES ON THE IMPLEMENTATION OF MONE'S NEW ASSESSMENT AND EVALUATION REGULATION TOWARDS ENGLISH LANGUAGE COURSES

Said Yasir KARAMAN
Çanakkale Onsekiz Mart University
School of Graduate Studies
Master of Sciences Thesis in English Language Teaching
Supervisor: Asst. Prof. Sedat BECEREN
24/11/2025, 120

This study examines the reflections and opinions of English language teachers and students in regard to the regulation implemented by Ministry of National Education in Türkiye. The regulation encompasses the assessment and evaluation of the English language which mandates that 50% of the course grade is based on speaking and listening exams. The regulation represents a shift toward communicative oriented assessment of the English language. The study aims to explore both students' and teachers' perceptions of the regulation and identify its advantages and disadvantages, thus revealing the challenges encountered by the stakeholders during its implementation.

A qualitative content analysis was employed as the research design to investigate the issue in depth. Data were collected from two groups which consists of 92 7th grade students and 7 English language teachers who were actively teaching at the same schools as the student group. Students completed open-ended written surveys, and teachers were interviewed using semi-structured interviews. The data were analyzed in MAXQDA software with guidance of literature on the subject.

Findings of the study suggested that both teachers and students recognized the benefits of the implementation such as speaking practice, improving communicative skills. However, particularly teachers expressed problems about the challenges encountered related to unclear assessment criteria, curricular overload, and limited instructional time. The students also voiced problems such as anxiety and stress caused by speaking exams and

academic concerns in addition to various difficulties they have encountered within the new examination system.

Based on the analysis of the collected data, the study discussed and evaluated how the new regulation was perceived by both teachers and students. The findings also provided detailed insights from the field, which contributed to drawing meaningful conclusions about the implementation of the policy.

Keywords: English Language Assessment, Assessment Reform, Educational Policy, Communicative Assessment, Speaking Evaluation

.

ÖZET

MİLLİ EĞİTİM BAKANLIĞININ İNGİLİZCE DERSLERİNE YÖNELİK YENİ ÖLÇME VE DEĞERLENDİRME YÖNERGESİ UYGULAMALARINA YÖNELİK ÖĞRETMEN VE ÖĞRENCİ DENEYİMLERİNİN İNCELENMESİ

Said Yasir KARAMAN
Çanakkale Onsekiz Mart Üniversitesi
Lisansüstü Eğitim Enstitüsü
Yabancı Diller Eğitimi Anabilim Dalı Yüksek Lisans Tezi
Danışman: Dr. Öğr. Üyesi Sedat BECEREN
24/11/2025, 120

Bu çalışma, Türkiye'de Millî Eğitim Bakanlığı tarafından uygulamaya konulan İngilizce dersi ölçme ve değerlendirme yönetmeliğine ilişkin olarak İngilizce öğretmenlerinin ve öğrencilerinin görüş ve yansımalarını incelemektedir. Yönetmelik, İngilizce ders notunun %50'sinin konuşma ve dinleme sınavlarından oluşmasını öngörmekte olup, İngilizce değerlendirmesinde iletişim temelli bir yaklaşıma geçişi temsil etmektedir. Bu araştırma, hem öğretmenlerin hem de öğrencilerin söz konusu yönetmeliğe dair algılarını ortaya koymayı, avantaj ve dezavantajlarını belirlemeyi ve uygulama sürecinde karşılaşılan zorlukları görünür kılmayı amaçlamaktadır.

Araştırma, konuyu derinlemesine incelemek üzere nitel içerik analizi deseniyle yürütülmüştür. Veriler, aynı devlet okullarında öğrenim gören 92 yedinci sınıf öğrencisinden ve bu okullarda görev yapan 7 İngilizce öğretmeninden toplanmıştır. Öğrenciler açık uçlu yazılı anketleri doldurmuş, öğretmenlerle ise yarı yapılandırılmış görüşmeler yapılmıştır. Veriler MAXQDA yazılımı aracılığıyla ve konuya ilişkin literatürden yararlanılarak analiz edilmiştir.

Araştırma bulguları, hem öğrencilerin hem de öğretmenlerin uygulamanın konuşma pratiği ve iletişimsel becerilerin gelişimi gibi bazı olumlu yönlerini fark ettiklerini göstermektedir. Ancak özellikle öğretmenler, değerlendirme kriterlerinin belirsizliği, müfredat yoğunluğu ve sınırlı ders süresi gibi çeşitli zorluklara dikkat çekmiştir. Öğrenciler

ise konuşma sınavlarının yarattığı kaygı, stres ve akademik baskı gibi duygusal tepkilerin yanı sıra sınav sistemine dair farklı güçlükleri de dile getirmiştir.

Toplanan verilerin analizine dayanarak bu çalışma, yeni yönetmeliğin öğretmenler ve öğrenciler tarafından nasıl algılandığını tartışmış ve değerlendirmiştir. Bulgular, sahadan elde edilen ayrıntılı içgörülerle birlikte, uygulamaya yönelik anlamlı sonuçların ortaya konmasına katkı sağlamıştır.

Anahtar Kelimeler: İngilizce Dil Değerlendirmesi, Konuşma ve Dinleme Sınavları, Ölçme ve Değerlendirme Yönetmeliği, Eğitim Politikası

TABLE OF CONTENTS

	Page
ETHICAL STATEMENT ……..……....	ii
ACKNOWLEDGEMENTS ……..	iii
ABSTRACT ……...	v
ÖZET ……..……..……..	vii
TABLE OF CONTENTS……	ix
THE LIST OF ABBREVIATIONS……..……...	xiii
THE LIST OF TABLES …….……	xiv
THE LIST OF FIGURES ……...…….……..	xv

CHAPTER ONE
INTRODUCTION

1.1.	Introduction...…….……....	1
1.2	Aim of the Study and Research Questions…….	2
1.3	Significance of the Study……..……..	3
1.4	Limitations of the Study…….……	4
1.5	Chapter Summary…….…….…….	5

CHAPTER TWO
LITERATURE REVIEW — 6

2.1.	Assessment and Evaluation in Language Education…..…….	6
	2.1.1. Assessment of Learning vs. Assessment for Learning.…….…….	7
2.2.	Theoretical Frameworks for Language Assessment and Evaluation……..	8
	2.2.1. Types of Assessment in EFL Teaching……...	8
	Summative and Formative Assessments ……	9
	Diagnostic and Placement Assessments ……	11
	2.2.2. Washback Effect…….……...	12
	2.2.3. Assessing Four Macro Skills……	13
	Assessment of Reading …….	14
	Assessment of Writing ……...	15
	Assessment of Speaking ……	15
	Assessment of Listening ..…….	18

2.2.4. Assessing ELF: Global Approaches……... 19
Overview of Global Approaches to ELF Assessment…… 20
Towards ELF-Informed Assessment Practices…….. 20
2.2.5. English Language Assessment and Evaluation in Türkiye…… 21
2.3. Communicative Language Teaching and Testing Framework…… 23
2.3.1. Historical Development of CLT…….…… 24
2.3.2. Principles of CLT…….……... 25
Communicative Competence ……. 25
Emphasis on Functional and Meaningful Communication……. 27
Learner Centered Approach…… 28
2.3.3. Communicative Language Testing …….……... 29
Key Features of Communicative Language Testing……... 29
2.3.4. Assessing Communicative Competence……. 30
2.3.5. MoNE Regulation Through Communicative Lens …… 32
2.4. Educational Policies and Regulations on Assessment and Evaluation…….. 33
2.4.1. Implementation Challenges of Policy-Driven Assessment Reforms…….. 34
2.4.2. The Context of Türkiye: Recent MoNE Reform…… 34
2.5. Chapter Summary…….…….. 36

CHAPTER THREE
METHODOLOGY 37

3.1. Research Design …….……... 37
3.2. Setting and Participants …….…… 38
3.3. Instruments and Data Collection Procedure ……..... 40
3.3.1. Student Surveys and Data Collection …… 40
Parental Consent …….……... 42
Language and Accessibility Considerations ……..... 42
3.3.2. Teacher Interviews and Data Collection……. 42
3.4. Data Analysis …….…… 43
3.4.1. Student Survey Analysis …… 43
Coding Process …….……. 44
MAXQDA Analysis Procedure ……. 45
Rationale for Theme and Code Development……. 46

Theme 1: Academic/Personal Development…….. 46
Theme 2: Skills Development…… 46
Theme 3: Challenges and Difficulties…… 47
Theme 4: Negative Attitudes and Experiences……... 48
Theme 5: Neutral/No Perceived Change…… 49
3.4.2. Teacher Interviews Data Analysis ……. 49
Coding Process ……. 51
MAXQDA Analysis Procedure ……. 53
Rationale for Theme and Code Development...……..... 55
Theme 1: Challenges and Difficulties Encountered……... 55
Theme 2: Perceived Advantages…… 56
Theme 3: Suggestions…… 57
Theme 4: Individual Solutions……... 58
3.4.3. Intercoder Reliability……. 59
3.4.4. Trustworthiness of the Study ……. 60
3.5. Chapter Summary.....……..…….....…… 61

CHAPTER FOUR
RESULTS 62

4.1. Introduction …….……..... 62
4.2. Research Question 1…..…….…… 64
4.2.1. Student Surveys ……..…… 64
Academic/Personal Progress ……. 65
Skills Development ……..……..... 66
Challenges and Difficulties Experienced ……..... 68
Negative attitudes/Experiences ……..……..... 69
Neutral/No Perceived Change ..…..…..…..…..…..…..…..…..…..…….... 70
4.2.2. Teacher Interviews ..…..…..…..…..…..…..…..…..…..…..…..…..... 72
Perceived Advantages ..…..…..…..…..…..…..…..…..…..…..…….... 72
Challenges and Difficulties Encountered ..…..…..…..…..…..…..…… 74
Individual Solutions ..…..…..…..…..…..…..…..…..…..…..…..…… 75
4.3. Research Question 2 ..…..…..…..…..…..…..…..…..…..…..…..…..…….. 78
4.3.1. Student Surveys ……..…..…..…..…..…..…..…..…..…..…..…..…… 78

Students' Perspective on the Benefits of the Regulation..…..…..…..……. 79
Students' View of the Limitations of the New Regulation ..…..…..…..…. 80
4.3.2. Teacher Interviews ..…..…..…..…..…..…..…..…..…..…..…..…..…..... 81
Teachers' Perspective on the Benefits of the New Regulation...…..…..…. 81
Teachers' View of Limitations of the New Regulation ..…..…..…..…..…. 82
4.4. Research Question 3 ..…..…..…..…..…..…..…..…..…..…..…..…..…..…… 84
4.4.1. Student Surveys……...…..…..…..…..…..…..…..…..…..…..…..…... 84
4.4.2. Teacher Interviews……..…..…..…..…..…..…..…..…..…..…..…..…. 85
Individual Solutions …..…..…..…..…..…..…..…..…..…..…..…..…..… 85
4.5. Chapter Summary...…..…..…..…..…..…..…..…..…..…..…..…..…..…… 86

CHAPTER FIVE
DISCUSSION, CONCLUSION, AND IMPLICATIONS 87

5.1. Introduction...…..…..…..…..…..…..…..…..…..…..…..…..…..…..…..…. 87
5.2. Teacher and Student Perceptions and Reflections of the Regulation.……. 87
5.2.1. Positive Perceptions and Reflections…..…..…..…..…..…..…..…..……. 90
Conceptual Agreement of the Implementation …..…..…..…..…..…..…... 92
Perceived Benefits: Skill Development and Academic Improvements….. 93
5.2.2. Negative and Mixed Perceptions and Affective Responses ……... 96
5.3. Reported Limitations and Implemented Solutions……. 97
5.3.1. Institutional and Pedagogical Barriers in Practice..…… 99
5.3.2. Adaptive Strategies and Solutions…..…..…..…..…..…..…..…..…..…… 100
5.4. Conclusion ..…..…..…..…..…..…..…..…..…..…..…..…..…..…..…..…..... 102
5.5. Implications and Suggestions …..…..…..…..…..…..…..…..…..…..…..…..…… 104
5.6. Chapter Summary …….. 106
REFERENCES...…… 108
APPENDICES……..……... I
APPENDIX 1 TEACHER INTERVIEW QUESTIONS FORM.……..... I
APPENDIX 2 STUDENT SURVEY QUESTIONS FORM.…….. III
APPENDIX 3 PARENTAL CONSENT FORM.……. V
APPENDIX 4 VOLUNTARY CONSENT FORM.……. VI
APPENDIX 5 ETHICS COMMITTEE APPROVAL……. VII

THE LIST OF ABBREVIATIONS

ELF	English as a Lingua Franca
ELT	English Language Teaching
EFL	English as a Foreign Language
CEFR	Common European Framework of Reference for Languages
MoNE	Mininstry of National Education
CLT	Communicative Language Teaching
CLTest	Communicative Language Testing

THE LIST OF TABLES

Table No	Name of Table	Page No
Table 1	Student Survey Themes	64
Table 2	Teacher Interview Themes	64

THE LIST OF FIGURES

Figure No	Name of Figure	Page
Figure 1	The dynamic of assessment, washback, and education	13
Figure 2	Student perception of academic and personal progress	66
Figure 3	Students' perceived skill development	68
Figure 4	Challenges and difficulties faced by students	69
Figure 5	Negative attitudes and experiences of students	70
Figure 6	No difference or difficulty experienced from student perspective	72
Figure 7	Teachers' perception of advantages	74
Figure 8	Challenges and difficulties encountered by teachers	75
Figure 9	Individual solutions implemented by the teachers	76
Figure 10	Suggestions proposed by the teachers	78
Figure 11	Benefits perceived by the students	79
Figure 12	Challenges and limitations students faced	80
Figure 13	Challenges and difficulties experienced by students	85
Figure 14	MAXQDA code co-occurrence map of Improved Grades	96

CHAPTER ONE
INTRODUCTION

1.1. Introduction

Assessment is an important part of English language teaching (ELT), through which learning can be controlled, progress measured, and decisions made. Whereas assessment previously emphasized testing knowledge about grammar, assessment is increasingly seen as evaluating students' ability to communicate in second language contexts. This reflects the growing emphasis in education on learner-centered, authentic, and competency-based instruction and assessment (Berry, 2008; Tsagari and Banerjee, 2022).

As a result of these international developments in language assessment, many language assessment policies in education systems have been changed to focus on communicative competence, the integration of skills and formative and performance-based assessments. Türkiye has recently moved toward stronger assessment practices through the introduction of the 2023 policy reform from the Ministry of National Education (MoNE). The policy states that speaking and listening exams now constitute 50% of students' final grades for English courses, a significant shift from the prior grammar-focused, summative assessment model. This change also reflects Türkiye's process to align with the Common European Framework of Reference for Languages (CEFR) and create more meaningful and functional uses of English in the classroom.

While the communicative assessment reforms are intended to be progressive, putting them into practice does present challenges. Teachers may find themselves faced with challenges relating to little institutional support, large class sizes, time constraints, and sometimes lack of exposure and training regarding alternative means of assessment (Coombe et al., 2017; Vogt and Tsagari, 2014) which may impact their ability to engage with the reform. The environmental challenges risk being greater in contexts like Türkiye where assessment has historically emphasized memorization, written accuracy, and centralized examinations.

As the MoNE regulation is recent, there is currently no empirical research that engages with how key stakeholders experience and engage with the reform, perspectives as

to teachers and pupils' experiences of their daily classroom interaction with the new assessments that ideally would enable them to experience communicative assessment. This study is intended to resonate with that knowledge gap and focuses on the perspectives and reflections of both teachers and students from public middle schools. Understanding how the reform takes shape, the challenges presented, how the reform impacts teaching and learning processes, will add to the body of literature for interpretation by policymakers, teacher educators, and researchers. Notably, the body of literature is starting to highlight the need for assessments that are context-sensitive, communicative, and pedagogically grounded, which are the foundations for this research, as outlined in detail in the next chapter.

1.2. Aim of the Study and Research Questions

This study aims to investigate the opinions, feelings, perceptions, and experiences of teachers and students with respect to the new English language assessment and evaluation regulations that MoNE implemented in 2023. A possible change from an emphasis on grammar to a more communicative approach is indicated by this updated system. In order to evaluate students' oral and listening comprehension abilities in the classroom and promote a more comprehensive development of language skills, the reform places a strong emphasis on the value of using authentic language.

In light of the aim of the study, the study addresses the following research questions:

1) What are the general reflections and perceptions of the teachers and students regarding the new regulation?

2) What are the perceived advantages and disadvantages of the new regulation?

3) What difficulties are experienced with the new regulation?

1.3. Significance of the Study

This study holds significance both in theoretical and practical categories. At the theoretical level, it adds to the growing body of research on English language learning and assessment. The shift in Türkiye' s English language assessment policy aligns with the modern approaches in language teaching that emphasize the communicative competence development. However, existing literature on communicative language acquisition and assessment mainly focused on Western contexts, and there is a lack of empirical research that investigates how such reforms and regulations are experienced and perceived in countries with different contexts (Kırkgöz, 2008). This research focuses on the important gap in the literature, by providing awareness and insights into how the MoNE regulation is implemented and experienced in schools in Türkiye.

At the practical level, this study presents valuable information for stakeholders and policymakers in the Turkish educational system. For policymakers at MoNE, the research provides input on how students and teachers are perceiving and experiencing the newly implemented regulation. The study gets feedback from both groups about what difficulties they have encountered, and what suggestions they may offer for improving its implementation and better implementation of new policies and regulations in the future. Such feedback is critical to establish that national reforms and policies on education fulfill their intended potential and outcomes with the least number of problems. Fullan (2007) claims that a successful educational reform relies not only on its design, but also on carefully taking into account the realities of the classroom practices, and the opinions of those who are directly involved.

For educators, the results provide valuable insight into the practical complexities and language learning struggles that exist for students based around the assessment of middle-school speaking and listening abilities. The thoughts and recommendations educators have about the challenges provide important information for the initiatives geared towards teacher preparation to use communicative assessment from the onset of classroom implementation. Furthermore, the research focuses on the practice of instruction and administrative support necessitated by the school for educators to be able to assess speaking and listening in a fair, effective, and optimally structured way.

For students, this study illustrates how changes in assessment practices and policies impact the learning opportunity, motivation, grades, and anxiety. In reviewing the positive reflections, and the negative experiences students navigate, this study provides the groundwork for ideas and future policies that reduce student anxiety, support skill-building, and provide a more constructive use of the assessment process as learning rather than simply measuring.

In summary, this study has broader relevance for educational investigation. It adds to the understanding of how worldwide educational agendas are adapted, and localized, in Türkiye. It adds to the evidentiary literature in relation to language assessment in Türkiye, and provides contextual knowledge of how various stakeholders notice, enact, and respond to national reforms at the classroom level.

1.4. Limitations of the Study

While this research provides useful insights into what student and teacher perceptions about the new English language assessment regulation in Türkiye, it also comes with important limitations. First, the research involved a small sample which consists of 92 7th-grade students and 7 English language teachers from three neighboring state schools in Çanakkale. While the participant group allowed for some rich qualitative and contextual information, these findings cannot be generalized to other regions, grade levels, and types of schools across the country.

Secondly, data was obtained from student surveys and teacher interviews. While this allowed for consideration of stakeholder experiences in great depth, it did not include classroom observations or official assessment materials, which could have provided an additional layer for triangulation. Furthermore, time constraints and limited access to institutions created restrictions that limited the possibility for more extensive longitudinal or cross-sectional studies.

Finally, an important limitation comes from the fact that the MoNE regulation is a recently instituted policy. There is no existing empirical research comparing this assessment model in English language classes with other similar studies in the country. Therefore, the

research of this practice in this study is an initial step to document the beginnings of such an experience, but the results need to be better researched and examined through further research.

In these ways, the study contained many undisclosed variables and so should be seen as an exploratory study. As such, repeating the research in a larger and more varied sample of different provinces, school levels, and time frames would help to make the applicability of the results stronger. Future research could also further improve data through the combination of a mixed-methods approach to make sense of how the regulation was interpreted and carried out by teachers in their classrooms.

1.5. Chapter Summary

In this chapter the context and rationale for the study have been presented in the context of the recent shift towards a more communicative and skills-based model in Türkiye's national English language assessment policy. The aim of the study has been outlined to explore how students and teachers experience and think about the new assessment regulation introduced by the Ministry of National Education in 2023. The research questions were developed to explore the experiences, perceived advantages and disadvantages of the regulation, and the challenges of the California of the policy. Their significance was emphasized in relation to the potential for the study to add to the empirical literature on language assessment in Türkiye and the implications it will have for future educational and policy decisions. Limitations around the scope of the study, sample size, methodological design and the recentness of the policy reform were also noted. The following chapter details the methodology that will be used to collect and analyze data in order to answer the research questions.

CHAPTER TWO
LITERATURE REVIEW

2.1. Assessment and Evaluation in Language Education

Brown (2004) points out that assessment is a significant part of language education which helps foster and utilize language development. In today's multicultural and globalized society, students' success in language learning depends greatly on effective assessment (Winna and Sabarun, 2023). In the ELT context, effective assessment plays a crucial role in supporting learners by motivating them toward meaningful learning which helps them reach their language goals. As a result, assessment and evaluation constitute a cornerstone of language teaching, as they provide guidance for both learners and teachers for monitoring progress and identifying areas that require scaffolding and support.

Taking language education policies into consideration, such as the one introduced by MoNE, and concerning the latest changes regarding assessment practices, it is important that practice is founded in sound principles. Assessment practices need to be shaped by reliable frameworks that are underpinned by principles of fairness, reliability, and communicative purpose. These frameworks, in language education and assessment provide an established basis to measure how assessments are designed, implemented, and evaluated for student learning outcomes.

The types of assessment can be classified based on their purpose (diagnostic, formative, or summative), their form (traditional or alternative), their nature (maximum or typical performance), and the interpretation of their results (absolute or relative) (Acar-Erdol and Yıldızlı, 2018). Understanding these distinctions is important for selecting assessment methods that align with the communicative aims of English language education.

In this context, assessment and evaluation should not only serve accountability but also act as a tool for guiding language learning. Rather than presenting disconnected perspectives, the literature demonstrates a common thread: assessment functions as both a measure of achievement and a driver of instruction. This dual role prepares the ground for later discussions of assessment of learning, assessment for learning, the washback effect, and the evaluation of specific language skills such as speaking and listening.

2.1.1. Assessment of Learning vs. Assessment for Learning

Assessment of learning refers to summative approaches that are traditionally used to track and measure learners' development in language education. These approaches often include standardized formats such as multiple-choice questions, cloze tests, and grammar-based tasks that assess discrete language skills like vocabulary, grammar, or reading (Rea-Dickins and Rixon, 1999; Brindley, 2010). While practical and easy to administer, these methods fall short in capturing students' meaningful and authentic communication.

In contrast, assessment for learning emphasizes formative practices that guide students during the learning process. This type of assessment is designed to support improvement by integrating tools that provide feedback, promote reflection, and connect assessment with authentic communicative use. With the recent changes imposed by MoNE, the need to balance summative traditions with formative approaches has become more evident in order to create a realistic representation of student performance.

Using data from formative assessments informs instructional processes and prepares students for summative assessments which enables teachers to create a supportive learning environment (Popham, 2008). Shepard (2000) also supports that the best assessment systems combine both approaches to promote meaningful learning.

Although summative methods are still beneficial and easy to administer and standardized, yet they often fail to measure communicative competence, which is at the heart of MoNE's updated English assessment policy. Since this study explores the experiences of teachers and students under the new regulation, it is essential to weigh the role of both summative and formative approaches. This comparison highlights the necessity of adapting assessment practices to national goals for enhancing English language learning in Türkiye.

2.2. Theoretical Frameworks for Language Assessment and Evaluation

Brown (2004) points out that assessment is a significant part of language education, which fosters language development rather than merely mobilizing and organizing it. In today's multicultural and globalized society, students' success in language learning depends greatly on effective assessment (Winna and Sabarun, 2023). In the ELT context, effective assessment plays a crucial role in supporting learners by motivating them toward meaningful learning and helping them reach their language goals. Theoretical frameworks for language assessment and evaluation are important for understanding what guides effective assessment practices. Considering current language education policies such as those introduced by MoNE, particularly the recent reforms related to assessment, it becomes necessary to refer to sound theoretical frameworks. These frameworks, informed by areas such as linguistics and psychometrics, provide a foundation for examining how assessments are designed, implemented, and evaluated in terms of their educatonal impact.

The types of assessment can be classified based on their purpose (diagnostic, formative, or summative), their form (traditional or alternative), their nature (maximum or typical performance), and the interpretation of their results (absolute or relative) (Acar-Erdol and Yıldızlı, 2018). Building on these categories, this section will examine the theoretical frameworks for language assessment in greater detail, particularly in the context of assessing listening and speaking skills, as well as English as a Lingua Franca (ELF), to better understand MoNE's recent policy regulation from the perspectives of both teachers and students. Understanding these frameworks ensures that assessment practices align with communicative competence and respond to the aims of the new regulation, with an emphasis on student learning and the broader Turkish education system.

2.2.1. Types of Assessment in EFL Teaching

Traditional assessment types have historically been used to track and assess learners' development within the field of language education. They typically consist of standardized formats, including multiple choice questions, cloze tests, and grammar-based assessment

tasks that assess language skills such as vocabulary, grammar or reading. While practical and simple to apply, traditional tests have the limitation of not being able to present students' meaningful forms of communication which reflect authentic language use. (Rea-Dickins and Rixon, 1999; Brindley, 2001). Taking into account the recent and new changes imposed by MoNE, we need to consider ways to adapt and/or integrate traditional tools to improve the capabilities for creating a more perspective, and realistic representation of student task performances and feedback.

To evaluate traditional assessments today, both their strengths and drawbacks must be considered. They provide structured and measurable insights into grammar and vocabulary, helping teachers identify specific language needs (Han and Kaya, 2014). However, when evaluating broader communication skills like speaking and writing, it is essential to see how vocabulary and grammar function in actual communication (Cheng et al., 2007). Therefore, while traditional tools still matter, they should be connected to communicative methods to better reflect students' real progress.

Traditional methods remain useful because they are easy to use, score, and standardize. Even so, they often fall short in capturing students' communicative competence, which is central to MoNE's updated English assessment policy. Since this study explores the experiences of teachers and students under the regulation, it is essential to weigh both the strengths and weaknesses of traditional assessments. This may help reveal whether new methods are needed to better match national goals for improving English language learning in Türkiye.

Summative and Formative Assessments

Summative assessment, which generally takes place at the conclusion of a unit or a course, generally aims to assess overall learning and assign grades (Scriven, 1967). On the other hand, a formative assessment is a progressive, and ongoing process of instruction meant to provide feedback, which is used to inform the instructional practices (Black and William, 1998). Understanding the difference between these two assessments is necessary

in order for educators to have a balanced and effective system which facilitates student growth and language learning.

Summative assessment is generally defined as assessment of learning that provides a summary of student achievements at a specific point. Its primary purpose is to evaluate learning outcomes according to the educational standards, assigning grades and certificates to indicate the level of students' proficiency (Gronlund, 2006). Summative assessments serve their purpose in the educational process by providing stakeholders (parents, employers, administrators, etc.) with valuable information about the knowledge and skills gained throughout the academic year, but they have their own negative consequences, including anxiety related to testing, prioritizing memorization, and restricted opportunities for feedback (Harlen and James, 1997). Thus, summative assessments should be developed to assess students, which is related to the student's learning goals and offer valid and reliable measures of language ability.

Formative assessment instruments may include quizzes, peer feedback, self-assessments, and class discussions. Studies on the effects of formative assessment suggest that effective formative assessment practices result in notable increases in student motivation, self-regulation and achievement (Black and William, 1998; Hattie and Timperley, 2007).

Considering the advantages of both, Shepard (2000) provides evidence supporting that the best assessment systems combine both assessments and present the opportunity to work best in student learning. Using data from formative assessments informs the instructional decisions and prepares the students for summative assessments and enables teachers to create a more supportive learning environment (Popham, 2008). Both summative and formative assessment has an important and different purpose in ELT. Summative assessments measure student accomplishment, while formative assessment affects learning process and teaching. Understanding the strengths and weaknesses of each assessment type and selecting the proper strategies is important to provide a balanced and enriched learning environment to support language proficiency and learning outcomes.

Diagnostic and Placement Assessments

Diagnostic assessments are a method of mapping learners' knowledge, strengths and weaknesses before instruction (Brown, 2016). Unlike general proficiency assessments, diagnostic assessments have a particular focus on certain language skill areas at a particular point in time. Teachers are able to take data gathered from diagnostic assessments to inform their instruction for intervention that specifically meets that given learners' needs, which is the most effective way to ensure that strategies used in teaching adhere to sound practice (Genesee and Upshur, 1996). This is also supported by Tomlinson (2017) who noted that diagnostic assessments, when used as feedback, will inform the instruction and personalized learning pathways which arguably, also offer learners support and guidance. Making diagnostic assessments is critical for EFL with contextually diverse learners. As indicated, this is especially relevant in EFL contexts where learners often have varying levels of English language exposure and using effective personalization with different learner backgrounds.

Placement assessments and tests are also a type of diagnostic assessments that place learners in appropriate groups according to the learners' current language competency (Bachman, 2004). The purpose in grouping groups of students is to avoid placing learners in inappropriate groups based on overly levelled learning material. In EFL, placement assessments are known to happen at the beginning of courses or programs of study to develop the most appropriate learning opportunity for learners or students. Alderson et al. (1995) argue that placement tests enhance learners' learning experiences because of developing a more homogenous and collective learning space, benefiting the learner as they progress at their level and receive appropriate instruction for their individual learning needs.

Ultimately, both diagnostic assessments and placement assessments serve to provide learning process with valuable information and opportunities to reflect upon in ELT and EFL contexts. These assessments are useful as they offer perception of the learner as in regard to strengths, weaknesses and overall learning development, collectively and over the time. Teachers are able to use information gleaned from both types of assessments to make informed decisions about their lesson, teaching progress and curriculum design, leading to more efficacious learning experiences.

2.2.2. Washback Effect

The term washback refers to the influence of an assessment on learning and teaching (Hughes, 2003). The influence of washback is important to EFL teachers, as it reveals how assessment can either support or limit effective second language instruction (Alderson and Wall, 1993). Washback includes the ways tests shape what teachers teach, how they teach it, and what students learn (Messick, 1996). Brown (2004) states that washback may influence curriculum design, changes in teaching materials, and student motivation. Generally, washback is discussed in terms of positive and negative effects. Positive washback occurs when assessment supports effective teaching and reflects the taught content and skills (Taylor, 2005). Wiggins (1998) describes this as an enriched curriculum when assessments align closely with learning outcomes. For example, teachers may use exam preparation to improve students' English by giving them a real-life reason to use the language (Watanabe, 2004). In contrast, negative washback occurs when assessment limits the teaching process and leads to restricted learning, often not supported by the teacher (Göktürk Sağlam, 2018). For instance, students may adopt a narrow view of language learning, equating success with test scores rather than actual communication or development (Shohamy, 2017).

In practice, the nature and strength of washback in EFL contexts depend on several factors. For example, formative, skill-based assessments designed to support CLT tend to emphasize meaning. Meanwhile, summative or discrete-point tests often push toward test-centered instruction and limit broader learning (Shohamy, 2017). Teacher beliefs also shape how assessments are used in practice. Green (2007) explains that teachers with communicative, learner-centered beliefs often apply assessment in more meaningful ways, while exam-focused teachers may teach only what the test covers, missing opportunities to develop communication skills. Additionally, the stakes of a test can intensify washback. High-stakes exams increase pressure on both teachers and learners to focus on scores and test-taking strategies. As a result, instruction may narrow and lose emphasis on language development (Alderson and Hamp-Lyons, 1996). Cheng (2005) argues that these factors interact within cultural, curricular, and institutional systems. Taken together, they show the complexity of washback in language education. A summary figure below, created in light of

the information given in this section, outlines these ideas and their connection to assessment in EFL.

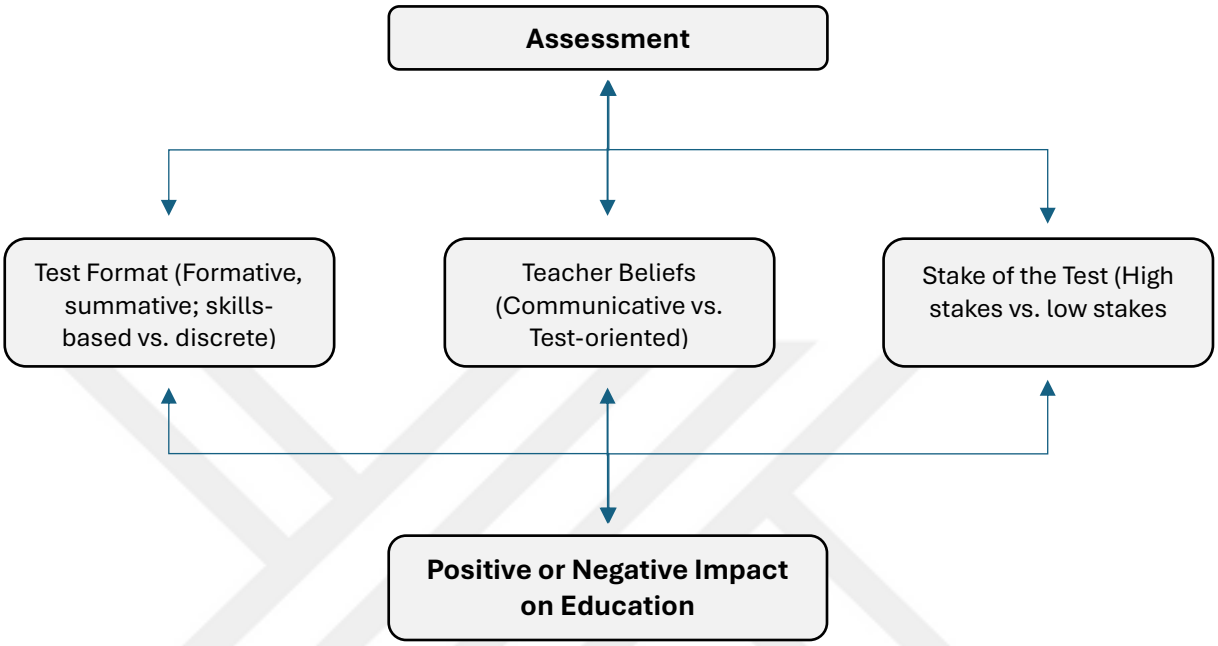

Figure 1. The dynamic process of assessment, washback, and education

2.2.3. Assessing Four Macro Skills

The assessment of four macro skills that are reading, writing, listening, and speaking is significant in the language evaluation (Brown and Abeywickrama, 2019). Each skill represents a distinct, yet interconnected aspect of communicative competence, and their assessment as integrated skills provides a holistic perspective on learners' language proficiency (Weir, 2005). In this respect, with the newly implemented regulation by MoNE, it is important to have a balanced and communicative approach to assessing all four skills. This section explores the principles and practices for assessing each of the four macro-skills by paying particular attention to the methods and difficulties of their assessment.

Assessment of Reading

Reading ability is a key skill in English as a Foreign Language (EFL), as it helps learners become familiar with written language, including various structures, vocabulary types, and discourse forms useful for language development. Reading supports the growth of comprehension skills such as recognizing grammatical patterns, understanding vocabulary in context, and identifying genre conventions. Reading assessments take many forms, including multiple-choice items, short-answer questions, cloze tasks, and matching activities. However, such assessments may not always reflect a full understanding of reading ability, as they often emphasize surface-level comprehension (Urquhart and Weir, 1998).

Reading assessments also vary depending on the text type and the cognitive processes involved. Grabe (2009) distinguished between lower-level reading skills like word recognition and syntactic parsing, and higher-order comprehension processes such as inference-making and text integration. Since reading is a receptive skill, assessment should go beyond isolated word or sentence understanding and focus on how learners derive meaning from entire texts. For instance, asking students to interpret an author's opinion, recognize implied meaning, or connect ideas across paragraphs demonstrate deeper comprehension.

In classroom settings, teachers usually create reading tasks that align with curriculum goals. Alderson (2000) emphasized the importance of context in reading assessments, noting that reading topics and texts should match learners' proficiency levels and interests. Effective reading assessments should also consider time constraints, familiarity with text types, and the authenticity of the reading material. When well-designed, reading assessments can enhance academic language development and help identify areas where students need further support (Fulcher and Davidson, 2007).

Assessment of Writing

The purpose of writing assessments in EFL is to assess students' ability to produce cohesive, organized, and meaningful written texts, and also the accuracy of their grammar, vocabulary, and proper organization of ideas when writing. Writing assessments can take many forms based on its purpose and type, for instance, guided writing tasks, portfolios, and self-assessment checklists. Each of these tasks provides information about student progress, and what changes and improvements are needed.

Scoring writing is often difficult because of the subjective nature of the skill. Analytical and holistic scoring rubrics are common methods of ensuring validity. Analytical rubrics assess writing by separating it into parts (e.g., content, organization, language use, etc.); holistic rubrics provide an overall score based on impression. Weigle (2002) observes that both methods have advantages and disadvantages, and the choice between methods can be made based on context and purpose.

Another issue related to writing assessment is related to genre influence and task familiarity since the learners may perform well on writing tasks involving genres or classroom assignments that they are familiar with. Hyland (2003) argues that genre-based approaches to writing assessment incorporate the need to teach students the structures and conventions of writing in each genre before assessing or responding to writing in that genre so that students are able to be assessed on their general language ability and their awareness of text conventions. Notably, a suitable writing assessment in EFL aligns with classroom objectives, assessment criteria that are transparently communicated to students, and meaningful feedback on student writing development provide meaningful feedback on writing skill development.

Assessment of Speaking

Evaluating speaking involves forming judgments and measurements about a language learner’s ability to use spoken language. The evaluation of speaking proficiency

involves unique challenges that require careful consideration which makes it arguably the most complex of all assessments compared to the other sub-skills (Luoma, 2004). Unlike other assessments in written form, speaking performance is often spontaneous, making it difficult to analyze and evaluate objectively. In addition, interviews might not be considered authentic to a real-life situation (Brown, 2010). In general, there are many ways to evaluate students' speaking skills, some are:

- **Oral Interviews:** Oral interviews are constructed or semi-structured interviews administered by the assessor with the learner (Fulcher, 2003). Interviews allow the assessor to evaluate the learner while they are speaking and evaluate a variety of speaking skills including fluency, accuracy, intonation, pronunciation, and use of vocabulary. The disadvantage of interviews is that they may take a long time and rely on subjectiveness of the assessor (O'Sullivan, 2000).

- **Presentations**: Presentations provide opportunities for students to exhibit their ability to express information in a clear and coherent fashion (Hughes, 2003). Presentations not only evaluate presentation skills such as body language and visual aids but also evaluate the student's linguistic ability.

- **Role-Plays and Simulations:** Underhill (1987) regards these types of assessments as simulations of real-life situations for communication, where learners must engage with other individuals who are assigned a specific role. Role plays are beneficial for assessing a wide range of speaking skills. For example, problem-solving, negotiation, and persuasion.

- **Discussions:** Discussions are a valuable means of assessing speaking and can be utilized in a classroom context to evaluate students' ability to articulate their thoughts, their listening accuracy and interaction with others. Weir (1993) has noted that rubrics must be developed to ensure coherent and fair evaluation practices in this form of assessment.

- **Technology Based Assessment:** The increased integration of technology in ELT has led to the use of online tools to assess speaking skills (Chapelle and Voss, 2011). Furthermore, Petersen (2022) has indicated that there are various forms of technology, such as recorded audio and video regarding speaking assessments, the

use of video or images as part of the questions, and peer-to-peer interactive assessments in video conferencing, are notable forms of technology that are used in speaking assessments.

In addition to the various methods and tools to assess speaking skills, there are also issues that must be taken into account. A major factor when assessing speaking in another language is that of speaking anxiety and its detrimental effects and influence on L2 learners' performance (Tercan and Dikilitaş, 2015). Many learners are apprehensive about speaking a foreign language, and in a classroom when the oral production of a language has the feel of a “test” rather than a chance to practice the language in a communicative setting as an example (Ansari, 2015; Zerey, 2008).

There are a number of reasons for perceiving anxiety while speaking, and it is possible to note some of these reasons as personal, social, and pedagogical causes. The main causes identified included low self-esteem, negative feedback, test anxiety, unpreparedness and lack of linguistic knowledge for assessment (Ansari, 2015; Şener, 2018; Tercan and Dikilitaş, 2015; Zerey, 2008). Öztürk and Gürbüz (2014) elaborated on the complex issue of speaking anxiety may arise from individual factors (unfamiliarity with EFL, fear of making mistakes, lack of self-confidence, fear of evaluation) environmental factors (negative classroom atmosphere, fear of peer pressure, and public speaking) and from school factors (limited practice opportunities, low quality instruction). This can potentially cause hurdles for learners as they try to develop communicative competence. Şener (2018) is very clear about the impairment having anxiety has on learners’ speaking performance and how it prevents them from communicating freely and building their speaking competencies or skills. Similarly, Ansari (2015) noted that speaking anxiety may hinder the adaptation to the learning environment and result in less achievement therefore, causing reluctance of speaking in the classroom, and unwillingness to participate in speaking activities.

In the context of this thesis examining 7th grade learners and the assessment of speaking in general, understanding the construct of speaking anxiety is important. In early adolescence, students generally have increased self-consciousness or self-awareness and have increased sensitivity to peer pressure, which makes them more susceptible to speaking anxiety (Saharani et al., 2023). By understanding the context of speaking anxiety, it is possible to better understand the student experience with the new regulation by MoNE.

Assessment of Listening

Listening is an essential component of EFL, as it allows students to understand spoken input and interact in speaking tasks. Listening includes the skill of interpreting and producing language, as learners must decode language in real-time, discern patterns in sounds, recognize meaning, and infer the message of the speaker. Listening assessments entail evaluating learners' ability to understand, interpret, and respond appropriately to spoken language. Listening assessment tasks may involve various methods including multiple-choice questions, comprehension questions, completing sentences, note-taking, or matching.

Listening comprehension is influenced by several factors beyond the text itself. As Buck (2001) points out, the listener's background knowledge, familiarity with the topic, and the cognitive effort involved in processing the input all play a role. For this reason, listening assessments should take these elements into account. Vandergrift (2007) similarly emphasizes that effective listening tasks incorporate both bottom-up processes, such as decoding sounds, and top-down processes, which rely on contextual clues or prior knowledge. Using both types of processing supports a more accurate understanding of learners' proficiency.

Rather than testing listening in isolation, classroom activities often integrate it into broader communicative tasks. Learners may listen to a recording and respond through speaking or writing, which increases their engagement and provides a more authentic experience. In addition to curriculum-based tasks, formative activities like audio journals or reflective listening logs can help students develop greater awareness of their listening strategies and track their progress over time (Field, 2009). Although listening assessments take place in different formats, these approaches give teachers a clearer sense of how students are growing as listeners in communicative contexts, while also helping to reduce test-related anxiety. In short, listening assessment should reflect the context in which it occurs, offers a variety of task types, and aim to capture the complexity of real-world listening.

2.2.4. Assessing ELF: Global Approaches

In the context of EFL, listening is one of the most important skills, allowing students to understand and respond to spoken English, which is one of the primary components of verbal communication in the classroom context and communication overall. It includes not just the process of hearing and requires the listener to make meaning at the moment, interpret the sounds, recognize the patterns, and anticipate what the speaker intends to communicate. Compared to reading, listening affords no time for self-review or slowing down, among other reasons that are implicated in learners' experience of processing information quickly enough. Listening tasks primarily focus on understanding spoken input, making a response to the spoken input, and making sense of an oral input. Some of the types of tasks, or activities, used for listening include multiple-choice, matching, note-taking, or filling in the blanks.

Comprehending spoken messages relies more on the meaning of the spoken words. According to Buck (2001), comprehension depends on prior knowledge, the kinds of topics students are familiar with, and the relevant cognitive effort required by the task. That's why listening assessment should take these elements into account. According to Vandergrift (2007), effective listening tasks involve the use of bottom-up skills, including hearing and recognizing sounds or words, as well as top-down strategies, such as using contextual and background information. Top-down and bottom-up strategies work together to give learners a better chance of improved comprehension of spoken language.

Listening is usually incorporated as part of a wider task rather than being assessed separately, learners might watch a video and respond with a short talk or write a summary. This keeps learners more engaged and gives the task more authenticity. In addition to using typical testing formats, formative tools, like listening diaries and audio journals, could be very effective for learners to reflect on their listening habits and track their development (Field, 2009). While listening assessments could be developed across different formats, these in-class assessment tools demonstrate on-the-ground what teachers are able to observe of their students' advancement in listening to and for real use.

Overview of Global Approaches to ELF Assessment

The use of English as a Lingua Franca (ELF) in education and the workplace is increasing globally, leading educators and policy makers to reconsider their earlier views on language assessment. In traditional contexts, assessment has often followed patterns based on native speaker events or has put heavy emphasis on grammatical correctness and pronunciation that conform to acceptable varieties or standards such as British or American English. In ELF contexts, the focus is on the degree of intelligibility and the successful resolution of interactions, even when disregarding native speaker expectations (Jenkins, 2009).

In light of tensions created by these changing priorities, a number of alternative assessment models have been proposed, such as Harding and Kremmel (2014), that incorporate interactional competence, flexibility, and strategic competence to assess in ELF contexts. These models raise challenges regarding how learners negotiate meaning, handle communication breakdowns, and adapt to multiple interlocutors. Similarly, Seidlhofer (2005) advocates for assessments that privilege pragmatic success over formal correctness through a focus on context and the pragmatic considerations at play in ELF communication.

While there is an increasing awareness of these normative issues, it might be a difficult task to apply this shift in perspective to large-scale standardized assessments because of emerging norms and the structured nature of assessment frameworks. As Rendle-Short (2006) argue, large-scale testing is not there to facilitate a monopoly on validity and reliability, qualities that may not align easily with ELF realities. Therefore, educational researchers and practitioners are globally reconsidering how to re-version tasks and rubrics to authentically capture the global realities of English use.

Towards ELF-Informed Assessment Practices

Developing assessment practices to exhibit ELF realities involves considering not only what is tested but also the way assessment is carried out. Cogo and Dewey (2012) mention that rather than focusing on native-like pronunciation or the strict application of

grammatical rules, ELF assessment seeks to assess intelligibility, pragmatic competence and the ability to manage interaction. Canagarajah (2006) claims that assessment practices should acknowledge the plurilingual repertoires of multilingual speakers and the strategies they use to communicate effectively. In this regard, rather than relying on traditional tests that focus on predetermined right or wrong answers, it is more effective to use assessment tasks like interactive dialogues, open-ended responses, or real-world simulation activities where multiple outcomes are possible.

Furthermore, ELF assessment recognizes a flexible and reflective approach to assessment. Given the training of teachers in ELF awareness, they are able to assess and interpret the performance of these learners, without relying on native speaker norms as a benchmark (Cogo and Dewey, 2012). To be aware of how communicative success comes from multiple accents, as well as diverse grammar and discourse styles that support mutual understanding, teachers must approach assessments with an open mind and look for evidence of competent communication.

In spite of some promising theoretical frameworks, the implementation of ELF assessments still has many challenges ahead. Roever and McNamara (2006) observe that the central tension in language testing lies in the relationship between reliability and authenticity, and assessment can be centered relatively between the two standards. Using ELF in classroom assessment, using ELF in formative feedback, and using ELF in self-assessments may serve as a positive step towards more equitable assessment systems. Ultimately, moving towards ELF assessment may align our examination practices with contemporary realities of communication, especially when English serves as a contact language among speakers of multiple, multilingual norms in classrooms.

2.2.5. English Language Assessment and Evaluation in Türkiye

The English Language assessment paradigm in Türkiye has mainly been developed around centralized testing and grammar-based methods. Traditionally, approaches to assessment predominately positioned language use in terms of knowing structures, and assessments were designed to measure only grammar or vocabulary knowledge (Kırkgöz,

2008). However, the past two decades has seen Turkish educational policy start to adopt communicative and skill-based approaches beyond grammar (Sönmez and Köksal, 2022). Notwithstanding this shift in policy, many teachers continued to exhibit preferences for more testing focused on instructional practices and shifted toward more communicative teaching methods and some lexically oriented assessment practices. This disparity between intended curriculum success and the actual execution of assessments continues to pose challenges in implementation.

In 2023, Ministry of National Education (MoNE) issued a revised assessment model for English language courses in secondary school. With the revised model, students can receive 50% of their final mark through either summative or formative assessment from productive skills - speaking and listening; and 50% from traditional written examinations (MoNE, 2023). This policy is now being positioned as an effort to support learners with their communicative competence while also considering assessments that align with the use of language in the world outside of LEEs. It represents the first high-level policy related notation that outlines not only what students know but also how they use that knowledge in practice.

The new assessment regulation provides benefits and challenges including opportunities for teachers to pedagogically utilize wider and diverse spectrum of interactive and learner-centred practices in their classrooms. However, a multitude of challenges exist in relation to class size and time constraints to enact and operationalize assessment mentioned, education, or training in assessment types, assessments themselves, and procedures, make it harder to apply the new system effectively (Demirpolat, 2015). The realization and implementation of actual communicative assessment into teachers' regular classroom practices is also impeded by structural base constraints. Overall, the newly amended policy is a positive first step toward communicative assessment in Türkiye, but further steps are needed in terms of teacher training, administrative support, and a shared system-wide understanding.

2.3. Communicative Language Teaching and Testing Framework

As English has spread globally, new communicative demands and expectations have emerged. Due to the expansion of globalization and digital technologies, English is now used extensively in several contexts (Canagarajah, 2006). Richards and Rogers (2001) point out that CLT is designed to help learners develop the ability to communicate in a second language. Krashen (1986) suggests that when learners are part of meaningful interactions, they naturally improve the skills they need to express their ideas. These types of experiences play a key role in building communicative competence.

The focus on communicative use in the MoNE regulation represents a move toward student-centered and communication-focused practices in language teaching and assessment. CLT supports meaningful language use and urges learners to use language effectively in real-world situations. In testing, CLT seeks to assess how learners perform in real communicative contexts, not just in terms of memorized words or grammar structures. This section explores whether the key principles of CLT align with the recent MoNE assessment policy. Clarifying this relationship may show whether the regulation enhances or restricts communicative development in Turkish EFL classrooms.

The goals stated in the MoNE regulation seem to reflect core aims of CLT. Exploring CLT alongside this regulation may provide useful insights. Previous studies have stressed the need to examine whether recent assessment systems truly promote communicative competence or still depend on grammar-based approaches. It remains important to ask whether the current system mirrors actual language use or remains focused on isolated language forms. This regulation's aim is to help students manage meaning, engage in spontaneous interaction, and take part in real conversations that are considered from a CLT and testing perspective. Using this framework could reveal how the policy supports or hinders communication-based English teaching in Türkiye.

2.3.1. Historical Development of CLT

Traditionally, the focus had been on identifying the "best" method for teaching languages. As far as the historical trajectory of language teaching is concerned, English Language Teaching (ELT) has a long history shaped by a number of teaching methods and approaches, including but not limited, Grammar-Translation Method, Direct Method, and Audiolingual Method. Each of these approaches had an instructional purpose. For instance, the purpose of the Grammar-Translation Method was to translate and teach grammar; the Direct Method avoided translation; and the Audiolingual Method focused on drills (Hunter and Smith, 2012). According to Hunter and Smith, this interpretation has been criticized for simplifying the diverse development of ELT, introducing imbalance, and treating instructional methods as overly fixed and unchanging. Early models often debated whether to emphasize grammatical instruction or promote conversational practice, highlighting a rule-oriented perspective on language learning.

In the 1970s, a major pedagogical shift took place with increased recognition of the learner and learning itself, rather than prioritizing the teacher or the act of teaching (Dinçay, 2010). During this era, the Cognitive Approach began to gain momentum. Although it was not yet widely established or precisely articulated, it began to exhibit more structure. The use of both deductive and inductive grammar approaches reflected a transition from a flexible orientation to a more prescriptive stance. During the late 1960s and early 1970s, we begin to see a solidification of Communicative Language Teaching (CLT)—drawing on ideas from psycholinguistics, sociolinguistics, and language acquisition. In CLT, Dell Hymes offered the idea of communicative competence (Ju, 2013). Hymes set up a view of proficiency that moved from grammar rules to practical, meaningful use of language.

Simultaneously, the notional-functional syllabus was introduced in teaching, advocating for teaching methods that focused on the use of language toward meaningful communication, intentional language use for tasks or purposes, and pragmatic speech acts (Hondo, 2018). The combined influence of communicative competence and the notional-functional syllabus allowed an emphasis of language teaching toward helping students express meaning (Pennycook, 1989). This emphasis of moving away from form-focused instruction into a purpose for communication, allowed teachers to focus on the use of

language in the classroom. Communicative approaches developed out of needs that were practical, and incorporated ideas from linguistics, psychology, and sociology. These ideas are seen today in language teaching as they emphasize learning language and communicating based on theory and research while adapting to the needs of classrooms and schools (Syarief, 2005). CLT developed as a flexible and learner-centered way to teach English through real communication.

2.3.2. Principles of CLT

Communicative Language Teaching (CLT) was developed as a response to a variety of form-focused approaches that were not able to highlight language as an instrument for communication in social and real-world terms. The reaction to form-focused approaches, however, was not merely to emphasize forms of grammatical correctness or isolated language skills, but rather to foster an ability to meaningfully and appropriately engage in real-world social interactions with others. This movement has been organized around a number of related principles around the way language is used socially. The principles of CLT recognize the functional use of language for both learners and speakers, highlight the importance of context, and foster the role of learners at the center of their own learning. The following sections detail the fundamental principles of CLT, starting with the foundation of communicative competence.

Communicative Competence

Communicative competence is an important concept to describe teaching and learning of languages which transitions focus from linguistic knowledge to linguistically based social language ability to effectively use language across various contexts. The concept of communicative competence was first introduced by Dell Hymes (1967, 1971) from his sociolinguistics works, serving as a corrective lens for what passed as linguistic theory (Cazden, 2011). Paulston (1974) referred to linguistic theories as being predominately focused on language competence based on internal cognitive abilities and knowledge of

grammatical rules that at the same time isolated language use from the realities of language in social contexts. Hymes (1972) defined communicative competence as a complex notion that encompassed dimensions of communication including knowledge of grammar, knowledge of language use in context, social norms, and the ability to act in context while using language. Vorwerg (2015) notes that Hymes characterized communicative competence as knowing how to use a language in social context in contrast to Chomsky's notion that focused only on grammatical correctness. Ma (2009) argues that Hymes' perspective of communicative competence was built on elements of grammatical competence and psycholinguistic, sociolinguistic and probabilistic dimensions of language embedded within the center of the Communicative Approach. Canale and Swain (1980) framed communicative competence in four category groups: grammatical competence, sociolinguistic competence, discourse competence, and strategic competence. Communicative competence is the ability to negotiate meaning, and is a vital term in CLT (Savignon, 1987). Ruiz (1985) agrees that a broad view of communicative competence impacted language teaching broadly, in that, it moved attention away from a focus on instructional methodologies that primarily addressed the teaching of grammatical forms. Communicative language teaching is concerned with learners developing functional language through reasoning and meaning embedded in communicative language activities (Littlewood, 2013). The focus of CLT is on developing reading, writing, listening, and speaking as the four skill channels of successful communication. CLT aims to support spontaneous, meaningful communication through classroom tasks involving information gaps, learner choice, and feedback for real communication experiences. CLT differs from traditional methods by emphasizing meaningful language use rather than isolated elements or rule-driven instruction (Sreehari, 2012). CLT promotes a shift from teacher-centered delivery to a more facilitative, learner-centered role in the classroom (Sreehari, 2012). This move forward could demonstrate a new focus on autonomy with opportunities for meaningful interaction in the language classroom atmosphere.

Communicative competence is crucial in relation to how language is used in the practical sense, especially spoken interactions and comprehension (Byram, 2021). Communicative competence refers to a broad understanding of the term, as it may be applicable to spoken and written forms of language. The point of communicative competence is emphasis on the ability to function in communicative environments, and for

this purpose might refer to communicating in real time (Rivers, 1973). Paulston (1974) argues that communicative competence includes the knowledge, skills, and disposition to engage in effective spoken interaction and convey meaning clearly in different situations.

CLT as a language learning and teaching method with a focus on communicative competence, attempts specifically to develop and nurture the ability to speak and the appropriateness of communication in addition to providing authentic tasks and autonomous activities for interaction that provides opportunities for competent use of a language. Therefore, speaking and listening in the language in this process is of fundamental importance to the definitions, development, and demonstrations of communicative competence.

Emphasis on Functional and Meaningful Communication

The aspiration of communicative language teaching is to develop communicative competence in learners based on the primacy of listening and speaking activities. From a communicative perspective, listening and speaking are both part of developing communicative competence in learners. The nature of communication and communicative competence, as defined in CLT, provides a foundation for designing communicative tasks and activities. The CLT approach to communicative competence allows for a shift away from structural approaches and verb tables, and towards functional communication that occurs in real-world settings. CLT is concerned with developing communication skills, especially speaking and listening skills (Mangaleswaran and Aziz, 2019).

The recent shift in MoNE's policy shares similarities with CLT in its emphasis on practical language skills. An ultimate aim of the CLT is to develop communicative competence that entails more than grammatical accuracy so that learners are able to accurately and effectively use the language in real-world settings (Al-Garni and Almuhammadi, 2019). According to Shevelova-Harkusha (2023), in regard to the relevance of grammatical rules, CLT provides the opportunity for learners to use the target language in a meaningful and interactive way. Communicative activities facilitate the engagement of

students through purposeful and meaningful interaction where there is an information exchange, expression of feelings, and expressed opinions (Rezalou and Yağız, 2023).

Celce-Murcia (2007) points out that communicative tasks and activities allow learners to develop competence in their speaking capabilities while developing their listening skills which helps learners listen and comprehend different types of speech featuring speech patterns and accents at the same time. Rezalou and Yağız (2023) further demonstrated, through the use of communicative activities, that EFL learners showed improved speaking performance, more positive perceptions of their speaking ability, and a greater willingness to participate. In Claudia Ho (2020)'s research, it showed that the undergraduates made considerable improvement in their speaking and listening competences after engaging in CLT activities.

Likewise, communicative activities may assist students in coping with speaking anxiety. Veenstra and Weaver (2022) found that doing communicative activities such as pair and group work reduced students' tendency to panic when speaking in English and their general anxiety when speaking. Together, these findings demonstrate how the emphasis on functional and meaningful communication in CLT leads to development of not only a linguistic ability, but also a sense of confidence and willingness to use the language.

Learner Centered Approach

The Communicative Language Teaching (CLT) framework places the language learner at the center, unlike traditional approaches that emphasize teacher control (Inamov and Holmatova, 2025). Compared to the Grammar-Translation Method, which is teacher-centered and limits student interaction to grammar rules, CLT emphasizes learner participation in classroom practice (Rouf and Sultana, 2015). CLT shifts the emphasis from the teacher as the central authority to recognizing the student as an active participant rather than a passive recipient. According to Jacobs and Farrell (2003), CLT's main aim is to develop communicative competence rather than isolated grammar knowledge.

This goal is supported by task-based, interactive activities that integrate all four language skills and offer students opportunities for authentic communication. Within CLT,

learners are expected to use language for various real purposes, negotiating meaning and conveying their message even if their grammar is not entirely accurate. Activities are designed to be student-centered, involve peer collaboration, and reflect purposeful, contextualized communication (Calderón Procel, 2024).

Allowing students to participate in these learning experiences shifts the teacher's role from the central authority of knowledge to the facilitator, who supports communication-based learning and monitors the learning. Communicative Language Teaching is designed to allow independence and autonomy for the students and to enhance meaningful learning and competence. From moving students from passive recipients to confident and communicatively competent learners ultimately contributes to the more comprehensive aim of developing and fostering independent and engaged language users.

2.3.3. Communicative Language Testing

According to the existing literature and the information mentioned above, communicative language teaching has revolutionized language education by shifting the focus from grammatical accuracy to communicative competence. As a result, language testing and assessment has also gone through significant transformations which lead to the emergence of Communicative Language Testing (CLTest). The principles of CLT had a profound impact on language testing. Communicative Language Testing aims to assess learners' ability to use the language in communicative situations effectively and appropriately (Bachman, 1990). This approach evaluates not only what learners know about the language, but also how well they can apply that knowledge in meaningful interaction.

Key Features of Communicative Language Testing

The tenets underlying the assessment of communicative competence in CLTest focus on testing language that happens in real communication contexts (Kaçar, 2023). It is possible to explain the model of CLTest based on a number of frameworks. These include:

- **Bachman's Model of Communicative Competence:** This model promotes contextualizing communicative competence utilizing a series of elements, including grammatical competence, textual competence, illocutionary competence, and sociolinguistic competence (Bachman, 1990; Bachman and Palmer, 1996).

- **Canale and Swain's Communicative Competence Framework:** This framework is based on Hymes' theory of communicative competence and defines communicative competence with respect to the three aspects of grammatical competence, sociolinguistic competence, and strategic competence, all of which are required to maximize effective language use (Canale and Swain, 1980; Canale, 1983).

- **Vygotsky's Sociocultural Theory:** Although not a theory for testing, Vygotsky (1978) acknowledged the importance of social interaction for language learning. Communicative Language Testing specifically includes collaborative tasks that occur in authentic settings, highlighting the sociocultural context of language use.

These theoretical foundations underpin the intentional assessment of the learners' ability to use language accurately, and effectively in real communication.

2.3.4. Assessing Communicative Competence

The focus in assessing communicative competence is to apply the principles and framework of CLTest to evaluate learners' ability to use language in real and functional situations. The CLTest's view of language highlights the capacity to communicate rather than treating language as a set of theoretical rules. Traditional language tests have long dominated language assessment, often relying on cognitive standards and focusing on measurable aspects of language knowledge. Although communicative language tests existed alongside these traditional models, they often still emphasized grammatical accuracy and cognitive benchmarks, while incorporating authentic tasks to measure communicative competence in both teaching and assessment (Liao et al., 2023).

Communicative language testing developed as a response to traditional testing models, which depended heavily on multiple-choice formats. These tests emphasized

reliability but often neglected authentic language use (Fulcher, 2000). Communicative language assessment focuses on real-world situations, making it a valuable component in evaluating language as part of learning progress.

Communicative competence includes several components: grammatical, sociolinguistic, discourse, and strategic competences, forming a complete understanding of communicative ability. Based on this framework, CLTest designs its assessments around values and standards rooted in communicative competence (Phan, 2008). Additionally, Bailey (1998, as cited in Phan, 2008) reports that a team from the Ontario Institute for Studies in Education outlined four main principles for designing CLTest assessments. The principles listed define the purpose and direction of the tests:

- **Start from somewhere:** The expectation of test takers in context. For example, what the test designers aim to assess, what they expect learners to do, and in which situations the target language should be used.

- **Concentration on content**: This principle stresses the need to focus on real-life context and tasks relevant to learners' needs. It emphasizes adapting content to learners' goals, interests, and proficiency levels.

- **Bias for the best**: Refers to creating supportive testing conditions that allow learners to perform at their best. Teachers are expected to provide necessary scaffolding so learners are not overwhelmed during testing.

- **Working on washback**: To ensure positive washback, test designers should create clear and consistent scoring criteria for both test takers and teachers. Aligning content with course objectives also contributes to beneficial washback.

In n summary, CLTest moves beyond traditional methods such as true/false and multiple-choice questions, which may test knowledge but fail to capture productive skills or communicative competence (Kaçar, 2023; Phan, 2008). CLTest offers a more holistic approach to assessment and aligns with communicative goals throughout the testing process. These principles are especially relevant to the present study, as they form the conceptual framework for evaluating how well the MoNE Regulation corresponds to communicative assessment principles in real classroom settings.

2.3.5. MoNE Regulation Through Communicative Lens

As alluded to in consideration of CLT and CLTest, MoNE's new regulation for assessment and evaluation presents a complex relationship between the theoretical intentions of the regulation and its practical realities. Understanding and analyzing the scope of a national regulation through internationally recognized theoretical frameworks and principles is an essential area of study (Shohamy, 2017; Burns and Richards, 2012). This analysis is not limited to communicative competence; it also involves evaluating whether assessment practices genuinely reflect the principles and values of CLT. In relation to competence from a communicative perspective, this raises the question of whether the regulation encourages teachers to implement practices that foster communicative competence or whether it brings little change. Understanding these issues through a communicative lens requires examining how the regulation is applied in real school contexts. These perspectives underline the importance of determining whether MoNE's new regulation supports or limits EFL students' development and use of communicative language competence in Türkiye. Rea-Dickins and Germaine (1992) argue that evaluating educational policy through its implementation allows practitioners to assess whether intended pedagogical values are being realized. This approach is especially relevant in Türkiye, where top-down policies often lead to varied interpretations and applications at the classroom level.

Despite its many strengths, the implementation of CLT and CLTest in practice also faces obstacles regarding the need for reliable assessment tools that produce consistent and valid results. Mansory (2020) notes that teachers may find it difficult to design assessment tasks that evaluate language use rather than focusing solely on grammar and accuracy. A study by Semiz and Odabaş (2016) found that assessment literacy among EFL teachers in Türkiye varies considerably, which may affect their alignment with the principles of CLT and CLTest and their ability to implement them effectively. Teacher development plays a key role in promoting meaningful change in communicative assessment practices (Scarino, 2013). Analyzing the MoNE regulation through the lens of CLT and CLTest allows for a clearer understanding of whether current assessment practices align with communicative competence principles. Evaluating the regulation's potential to improve English language teaching and assessment for EFL learners requires grounding in CLT, CLTest, and related teacher development literature. Ultimately, this review aims to support informed decisions

in teacher education, curriculum development, and policymaking to promote more communicative approaches in Türkiye's EFL instruction.

2.4. Educational Policies and Regulations on Assessment and Evaluation Practices

The role of policy and regulation is becoming an increasingly significant aspect of reshaping the environment of language assessment and evaluation in educational systems worldwide. Educational policy and regulation shape the educational system; they define not only the curriculum and instructional practices but also how assessment is carried out (Shohamy, 2017; Fulcher, 2000). Gipps (2012) argues that decisions about policy indicate the skills and knowledge that are prioritized, as well as the standards for assessing and the manner of assessing the skills and knowledge taught. That is, assessment practices reflect policy intentions and play a key role in how policies are implemented in schools. Globally, frameworks such as the Common European Framework of Reference for Languages (CEFR) is an example of how policy can shape assessment practice across diverse contexts (McNamara, 2017). The CEFR has been widely adopted, leading to standardization in many contexts, often without taking into account different pedagogical traditions or learner diversity. McNamara even argues that policy initiatives in this respect may risk imposing policies and constructs of education that do not include the cultural diversity of language learning.

Policy reforms that aim to modernize assessment practices often result in a gap between the policy's intentions and what happens in classrooms (Looney, 2011). In a number of educational contexts, Vogt and Tsagari (2014) maintain that teachers may lack the assessment literacy, institutional support, or resources needed to implement the goals of the policy effectively. Accordingly, policy decisions and regulations in assessment can lead to uneven or unintended outcomes in educational systems, especially in classroom practices of teachers and learners.

2.4.1. Implementation Challenges of Policy-Driven Assessment Reforms

There is an ongoing worldwide transition from grammar-oriented approaches toward communicative language teaching, which has implications for assessment reform in many contexts (Richards, 2006; Brown and Abeywickrama, 2010). In communicative assessment, the goal of assessment is to evaluate a learner's ability to communicate in authentic interaction, which means moving away from conventional grammar tests to performance-based assessments based on speaking and listening (Bachman and Palmer, 1996; McNamara, 2017).

However, there still remain significant challenges to successfully implementing communicative assessment within the instructional context. Green (2007) found that teachers grappled with how to fully realize communicative assessment due to their limited time, large class sizes, and a packed curriculum. The introduction of national standardized assessments may intensify these challenges and produce a negative washback, even when reforms stem from valid goals.

In addition, Birenbaum et al. (2015) observe that while approaches in formative assessment, typically known as Assessment for Learning approaches, are drawing attention globally, they found many challenges with classroom implementations due to the tension caused by fixed summative assessments. This tension creates a larger challenge for matching aims of policies and regulations with what actually happens in classroom reality.

2.4.2. The Context of Türkiye: Recent MoNE Reform

The practice of language assessment in Türkiye has typically inherited a grammar-focused tradition that emphasizes written tests, vocabulary development, and grammatical accuracy (Kırkgöz, 2008). These traditions have influenced the practices of teachers and students so that language is conceptualized as an accumulation of language rules and form, rather than an instrument to communicate meaning.

At earlier times, languages, such as Arabic, French, and German, served social and political contexts; however, the English language emerged as the primary foreign language

in Turkish schools (Sönmez and Köksal, 2022). During this Republican period, classroom teaching and curriculum design prioritized grammar-based assessments and teacher training (Nergis, 2011). In terms of teacher competency, the Gazi Institute and other faculties of education were responsible for training teachers, yet the training programs often promoted linguistic knowledge over pedagogical and assessment literacy (Mahalingappa and Polat, 2013). Tokçalar (2025) described language education policies throughout the Ottoman period to the present time and described how each period implemented policy reforms, yet teacher preparation often failed to recognize changes in classroom demands and communicative/efficient standards and, ultimately, upheld the previous exam system.

Recently, Türkiye has witnessed educational reforms aimed at creating educational experiences that mirror CEFR standards as well as the importance of English as a lingua franca (Sönmez and Köksal, 2022). Unfortunately, persistent issues of large classroom sizes, curricular restrictions, and a lack of professional-development have hampered the purposeful integration of change in classrooms (Paker, 2012, pp. 139–149). The traditional form of exams, which has focused heavily on form rather than meaning, memorization, and grammatical accuracy, has historically inhibited the establishment of communicative competence within education in Türkiye.

The 2023 reform from the MoNE could be recognized as a transformative policy shift that intends to emphasize globally recognized communicative language education within the context of Turkey. In this reform, speaking and listening assessments are weighed as fifty percent of a student's final evaluation (MoNE, 2023). This shift represents Türkiye's move toward communicative assessment practices based on CEFR. Nonetheless, it is important to recognize that without improvement of teacher education, professional development and material, policies will not be enough (Mahalingappa and Polat, 2013; Oktay, 2015). If the policy does not explicitly connect intentions with teacher knowledge and classroom practice, teachers are left to implement policies at a surface level rather than an effective practice under them (Das et al, 2014).

To summarize, the Turkish case highlights the intricate and multi-faceted relationships between educational history, policy history and reform, assessment practices, and teacher education. Although the new MoNE regulation has the potential to shift towards

a more communicative and skill-based assessment practice, its success depends on addressing the structural and institutional issues raised in literature.

2.5. Chapter Summary

This chapter illustrates key concepts and studies that formed the foundation of the present research. The chapter opened with concepts of English as a Lingua Franca (ELF) and its historical context in Türkiye. The chapter then moved on to Communicative Language Teaching (CLT) and discussed the types of tests that support these approaches in addressing real-life communication needs of learners who engage with English beyond the classroom.

Communicative Language Testing was explored in depth and the chapter examined how Türkiye's new regulation from MoNE aligns with this perspective. The chapter also presented the various models of assessment, types of tests, washback, and the assessment of four key language skills. There was also a notable focus on how global frameworks are addressing ELF and how such global terms have begun to influence classroom assessment.

Finally, the chapter discussed how national policy connects with actual assessment practices in schools. It also reviewed both past and current regulations in Türkiye, with a focus on the recent reform that gives more weight to speaking and listening exams. The chapter ended by addressing key challenges in practice, such as time limitations, unclear rubrics, and misalignment with the national curriculum. Overall, these sections provided the foundation needed to understand the purpose and relevance of the current study.

CHAPTER THREE
METHODOLOGY

3.1. Research Design

This study employed a qualitative research design to explore the experiences and perceptions of seventh-grade students regarding the new English language assessment implemented by the Turkish Ministry of National Education (MoNE). The qualitative research design was suitable for this study as it enables researchers to examine complex social phenomena through the lens of participants (Creswell and Poth, 2018). Qualitative inquiry differs from quantitative approaches, such as statistics, which emphasize the measurement of numbers and inferring meaning from those numbers. The foundation of qualitative research is to understand how individuals arrive at conclusions regarding the internal meaning ascribed to their experiences and how they interpret them within their context (Denzin and Lincoln, 2011). With this study focusing on students' subjective reflections regarding a new assessment policy, qualitative inquiry provided an opportune space for collecting rich, nuanced, contextual data.

The foundation of this research design is qualitative content analysis, a systematic process for condensing large amounts of textual data into fewer content categories, grounded in explicit coding rules (Schreier, 2012). Content analysis allows a researcher to discover patterns, themes, and categories across participants' responses, and interpret those findings with respect to the research questions. This method is particularly important for educational research, as participant, open-ended feedback can vary greatly in form, tone, and depth (Hsieh and Shannon, 2005). This research used an inductive approach to content analysis: themes and categories were derived from the data rather than pre-identified, although the researchers developed initial codes using the related literature and advisor input. By exploring students' perceptions and reflections on the new English assessment system, the research design acknowledged the context of experience, individual differences, and the individual process of meaning making that occurs in educational settings.

The research questions in this study were developed to provide more holistic insights into students' experiences in the classroom, including the innovations they observed, the skills they built, and the barriers that arose when working within the new English assessment

system. Qualitative designs provided a flexible approach that allowed for both commonalities and individual variations in students' experiences to be captured within the sample. This was done to ensure that the findings represent the full diversity of perceptions within the sample.

3.2. Setting and Participants

The research took place in three public middle schools in the province of Çanakkale in Turkey. Çanakkale is a socio-culturally diverse environment in which publicly funded schools cater to students from diverse socio-economic conditions. These schools were chosen because they represent public school contexts where the new assessment policy developed by the Ministry of National Education has been implemented throughout the country. Therefore, the research findings would have some local relevance and be based in a typical public-school context. The public middle schools in Türkiye reflect a centralized education system where curricular and assessment changes are mandated nationally for all schools. Having both students and teacher participants in this study enabled a rich and in-depth exploration of how the new MoNE assessment policy affected learning experiences and classroom practice. Students were able to provide insights into how the policy related to their learning processes, classroom experiences, and skill development while teachers contributed professional insights related to pedagogical changes, classroom challenges, and implementation experiences.

In this study, purposive sampling was employed, targeting participants who had a direct involvement and experience in the application of the new English language assessment regulation introduced by the Turkish Ministry of National Education (MoNE). Purposive sampling allows researchers to intentionally choose individuals who possess relevant knowledge and experience regarding the phenomenon under investigation. In this study, English language teachers and 7th-grade students were selected because they represent the primary stakeholders affected by the recent assessment changes, making their insights essential for addressing the research questions.

Participants were divided into two distinct groups, a participant group of students and a participant group of teachers. The student group consisted of 93 seventh-grade students who were engaged in English language at the schools during the 2024–2025 school year. Student age ranged from 12 to 13 years, and they represented a diversity of academic achievements and backgrounds. This provided a broader landscape of how students of various profiles experienced the introduction of the new assessment policies that were enacted in the public schools in Turkey.

Along with the student participants, there were seven English language teachers that participated in the research. All of the teachers were also reportedly teaching seventh-grade classes at the same time as the research and were purposely selected to align with the students' sample. The criterion on which these teachers were selected was that they were directly delivering English ideally to the same students who responded to the survey. These criteria ensured alignment between the participants and allowed for an exploration of the extent to which teachers understood and implemented the assessment reforms that their students were experiencing.

Prior to data collection, formal research permission was obtained from the Turkish Ministry of National Education from its online research permission portal at (https://arastirmaizinleri.meb.gov.tr/). This ensured that the research complied with the governmental policies and regulations governing research in public schools.

Parental consent was secured for all student participants utilizing a formal consent form that outlined the study's purpose, emphasized the voluntary nature of participation, and highlighted measures taken to protect confidentiality and anonymity. Only the students whose parents granted written consent were included in the study. When the surveys were delivered to the students, they were orally informed about the study's purpose, their right not to participate or to withdraw from participation without drawbacks, with guarantees of anonymity.

For teacher participants, individual informed consent was obtained verbally and in writing before the interviews commenced. Teachers received explanations about the study's purpose, process and how they would be treated as participants and the intended use of their input. Interviews were scheduled at the teachers' convenience during non-teaching periods and consulted and coordinated with school administrators so that data collection did not

disrupt the school's schedule and practices. All interviews were audio recorded with participant permission.

Participation was completely voluntary for the student and teacher participants. Participants were clearly told that they could withdraw at any time and that no action would be taken against them for withdrawing. Participants were also assured that their data would be kept confidential and all identifying information would be removed during the research and reporting of findings.

3.3. Instruments and Data Collection Procedure

The instruments used and the procedure of data collection followed are given under two categories for each group of participants. Firstly, student surveys data collection steps are explained, sequentially teacher data collection procedures are presented.

3.3.1. Student Surveys and Data Collection

The primary instrument for data collection from the student component of this study was an open-ended survey that aimed to capture students' reflections on their classroom experiences through the new policy of assessing their English language. The survey was designed based on reading research in language testing, assessing communicative competence, and assessment washback (e.g., Brown and Abeywickrama, 2010; Shohamy, 2017). Draft questions were originally written to align with theorized frameworks and then consulted on with three English Language Teaching (ELT) experts to establish clarity, content validity, and appropriate age marker.

The survey underwent consultation by practicing teachers to ensure the questions were structured and coined at a level suitable for seventh graders. Cultural suitability and contextual aspects of the survey were also accounted for to make certain the Turkish middle school students would be able to engage in a way that produced informative data.

The final version of the survey consisted of three open-ended questions:

1. “How did the new assessment system affect your English learning process in the classroom? If there were any new practices you experienced, could you give examples?”

2. “How did the new assessment and evaluation system help you use your English speaking and listening skills, and what kind of contribution did it make to these skills?”

3. “What difficulties did you face under the new assessment and evaluation system?”

The intention of these questions was to elicit detailed, descriptive accounts of students' classroom experiences, understandings of innovation, skill development, and the barriers experienced when transitioning to the new assessment model.

Data collection was prepared and conducted in a way that would not interfere with regular school activities in a manner with great ethical rigor. Before distributing the surveys to students, the researcher met with classroom teachers to explain the study and the general intent and context of the data gathering process. This also allowed for the opportunity to collaboratively select a time (during the normal English language lesson hours) that would be suitable to the classroom teachers and appropriate to the researchers’ timelines for the survey and the participant’s timelines.

Prior to administering the surveys, students had a verbal briefing about the study and the survey. The briefing presented an outline of the study’s intentions, assurances of confidentiality and anonymity, emphasized that participation was voluntary, assured students that there were no right or wrong answers and that the researchers' interest was in students' honest opinions and thought process. Students were encouraged to communicate if they had concerns about the study and that they were welcome to ask any questions related to the survey and research study. Most of the students required little instruction or clarification, however, additional clarification was provided to any of the students who desired further understanding.

The surveys were carried out within the regular class schedules, during the first 15 minutes of a class to ensure minimal disruption to the planned activities. Teachers were

present but did not intervene to allow students to write what they wanted without pressure. Surveys were collected anonymously, and no identifying information was targeted or recorded as a way to protect the survey participants and their opinions.

Parental Consent

As a part of ethical data collection, parental consent was a mandatory requirement which meant it had to occur before the student participated in the study. The parental consent form provided to parents explained the intent of the study, what would be expected of the student if they chose to participate, that participation was voluntary, and brief explanation of how the data would be kept confidential and used for academic purposes. The form also explained that the surveys would be completed in regular class time and that no sensitive personal information would be collected. Only students whose parents filled out and signed the consent forms were invited to participate in the survey portion of the study in compliance with ethical policies for research that involved minors and young students.

Language and Accessibility Considerations

The surveys were developed and delivered in Turkish, the respondents' language, to optimize understanding and facilitate students fully expressing their thoughts. It also eliminated language proficiency as a barrier to participation and allowed students to reflect on their classroom experience comfortably and authentically.

3.3.2. Teacher Interviews and Data Collection

In addition to the student surveys, semi-structured interviews were conducted with seven teachers of the English language who were actively teaching seventh grade in the schools participating in the study. The purpose of the interviews was to elicit teachers'

reflections on the implementation of the new evaluation policy, teachers' experiences administering the revised exams, and teachers' observations regarding adaptations for students and changes to classroom practices.

Before interviewing the teachers, the teachers were provided with an explanation of the purpose of the study, the methods used to conduct the study, and the ethics surrounding their participation. Each teacher provided informed consent through a written consent form that described the voluntary nature of their participation and responses, the confidential nature of their responses, and the secure way in which the audio recordings of the interviews would be conducted. It was outlined that the distinct possibility for teachers to participate on a voluntary basis, that an audio recording would be made for accuracy, and that recordings would be securely held in a way unavailable to anyone, except the researcher, for the duration of the data collection phase of the research.

Interviews were conducted at times that were less disruptive to teachers' professional responsibilities. Interviews took place during breaks or release time within unoccupied classrooms, which provided a quiet and private space for discussion. Interviews were scheduled to avoid teachers' class times; accordingly, interviews were brief, typically 10 minutes in length. While the teachers' time was valued by keeping the interviews brief, rich qualitative data was gathered.

3.4. Data Analysis

Data analysis process is presented in two sections as it includes two groups as students and teachers.

3.4.1. Student Survey Analysis

The data analysis in this study used a qualitative content analysis framework to capture the complexity, diversity, and richness of student reflections about a new assessment system. This assessment system was new because of a policy shift from Turkish Ministry of National Education (MoNE). Qualitative content analysis was preferred over quantitative content analysis because of its systematic and flexible, and replicable aspect of categorizing

qualitative content (Schreier, 2012). Qualitative content analysis is better suited for exploring student-defined subjective experiences or perspectives through surveys. While quantitative content analysis relies on a direct counting of reoccurring words and phrases, qualitative content analysis values the conceptualization and interpretation of themes and meanings (Hsieh and Shannon, 2005). This is what Braun and Clarke (2006) refer to when they describe qualitative content analysis as aligning with the interpretivist paradigm. The is the reason for this study's approach, allowing researchers to understand the ways in which individuals make meaning based in a particular social context-specific group of seventh-grade students revisiting their own considerations about the English language assessment system established by the Ministry of National Education (MoNE).

Since a qualitative content analysis frame was used in our study, it was possible to identify and describe patterns in students' open-ended comments on their surveys about changes in their classrooms, development of skills, and challenges they met under the new policy. Inductive qualitative content analysis allows categories to emerge from the data while still being aware of the literature in the field of study and expert input, thus giving shape but not confining the findings to predetermined categories.

Coding Process

The coding process started by reading all the students' responses to become familiar with the range and texture of the data. Each student's written response counted as a single unit of analysis; an individual portion (sentence or phrases) counted as a secondary unit of coding when there were multiple ideas expressed in one written response. Each initial code was derived from the recurring concepts in the students' comments and were refined through constant comparison of codes hoping for a level of consistency and conceptual clarity across the data set. Throughout the coding process, codes were developed not only to give labels for patterns noticed in the data, but they were also informed by theoretical concepts and prior experiences with washback effects, language assessment reform and communicative language teaching (Shohamy, 2017; Brown and Abeywickrama, 2010; Richards, 2006). To help add validity and analytical rigor into the code system two English Language Teaching (ELT) experts were asked to review the code list for appropriateness, clarity, and

completeness. This expert validation process provided a level of accountability and assurance that the codes were reflective of the realities of the Turkish classroom context while giving consideration towards the theoretical constructs from the literature.

MAXQDA Analysis Procedure

Data analysis for this study was facilitated using MAXQDA (the computer-assisted qualitative data analysis software). MAXQDA enabled systematic organization, coding, and visualization of the data, and the analysis steps summarized for the analysis process are:

Importing data: The student responses were stored as an Excel spreadsheet, which were imported into MAXQDA through the "Import Spreadsheet" function. Each student's response from the survey question was organized as independent variables (Q1, Q2, Q3) with a participant ID to aid in efficient organization.

Creating a hierarchical code system: In MAXQDA's Code System window, a hierarchical code system was arranged to demonstrate main themes (parent codes) and more focused related subthemes (child codes). The coding structure allowed for broad themes to be arranged with subsequent meanings produced throughout the analysis of the data.

Visualization and verification: The software's visualization tools, including the Code Matrix Browser, Code Frequency Table, and Word Cloud generator, were then used to interrogate coding patterns or co-occurrences, creating visual representations of the dominant and subsidiary themes. This process verified coding consistency and provided descriptive summaries of the results.

Iterative refinement: The researcher used iterative refinement of the coding scheme as the research moved through the dataset by adding emerging codes as themes were conceptualized and restructuring redundant codes to provide analytic clarity. Such iterative analysis is consistent with standards for qualitative research and content analysis.

Rationale for Theme and Code Development

The final coding system was created around five primary themes, each with individual subcodes that indicated particular patterns amongst the student data. The following provides a description of each theme, its associated codes, and analytic definitions:

Theme 1: Academic/Personal Development

This theme illustrates how students perceived improvements in academic outcomes or personal learning progress related to the new assessment system.

- **Improvement in Learning:** General perception of improved learning with the new system.
- **Classroom Participation:** Increased active engagement during lessons.
- **Improved Understanding:** Enhanced comprehension of English content.
- **Better Learning:** Perception of more effective learning.
- **Feeling of Improvement:** Subjective view of getting better at English.
- **Improved Grades:** Perception of positive impact on exam grades.
- **Increased Motivation:** Increased enthusiasm or willingness to study English.
- **Enhanced Self-Study Habits:** Improvement in self-study practices.

Theme 2: Skills Development

This theme reflected perceived improvements in skill development, particularly in speaking and listening skills.

- **Speaking Skill Improvement:** Overall improvement in speaking skill.

- *Improved Fluency*: Speaking more easily and coherently.
- *Speaking Outside the Classroom*: Using English in situations outside of school.
- *General Improvement*: Non-specific perception of improvement in speaking.
- *Confidence in Speaking*: Feeling more confident in speaking English.
- *Pronunciation Improvement*: Perceived improvement in articulation and pronunciation.

- **Listening Skill Improvement:** Overall gains in listening skill.
  - *General Improvement*: Non-specific sense of listening improvement.
  - *Word Familiarity*: Recognizing more words while listening.
  - *Listening Comprehension*: Understanding spoken English better.
- **Memorization and Retention:** Being able to memorize and remember vocabulary or grammar content.

Theme 3: Challenges and Difficulties

This theme captured obstacles, barriers, and adaptation problems students encountered with the new system.

- **Assessment-Related Challenges:** Difficulties related to the new assessment structure.
  - *Time Management*: Having trouble managing time during assessments.
  - *Exam Format*: Difficulty understanding or adjusting to new exam structures.
  - *Scoring Confusion*: Uncertainty about grading criteria or scoring.

- *Exam Overload*: Feeling overwhelmed by the excessive number of assessments.

- **Difficulty in Adaptation:** Difficulty adjusting to the new system in general.
  - *Initial Adaptation*: Struggling at the early stage when the system was first introduced.
  - *Examination Mismatch*: Perception that tests did not match content or centralized exams.
  - *Lack of Practice*: Feeling of having limited opportunities for practice.
- **Difficulty with Listening Tasks:** Problems related to listening exams.
  - *Technical Issues*: Problems during listening tasks regarding technical audio issues.
  - *General Difficulty*: Non-specific difficulty related to listening exams.
  - *Fast/Unclear Speech*: Trouble understanding rapid or unclear audio regarding the listening track.
- **Difficulty with Speaking Tasks:** Challenges related to speaking exams.
  - *Recalling Words*: Difficulty retrieving necessary words while speaking.
  - *General Difficulty*: General difficulties with speaking exams.
  - *Pronunciation Difficulty*: Difficulty pronouncing words correctly.
  - *Performance Difficulty*: Issues with delivery during speaking.

Theme 4: Negative Attitudes and Experiences

This theme captured the negative emotional and psychological reactions and negative attitudes expressed by students about the new system.

- **Inhibited Learning:** Feeling that learning was hindered.
- **Decreased Grades:** The belief that the grades were lower due to the new system.
- **Anxiety and Stress:** General anxiety or stress related to the reform.
    - *Fear of Making Mistakes*: Anxiety about making mistakes in exams.
    - *Speaking Anxiety*: Nervousness and stress when speaking English.
    - *Academic Anxiety*: General academic stress.
    - *Test-Related Anxiety*: Anxiety that occurred before or during exams.
- **No Improvement Felt:** The belief that no learning progress was achieved.
- **Loss of Motivation:** Reduced motivation to learn under the new policy.

Theme 5: Neutral/No Perceived Change

This theme reflected students' responses claiming that there were no perceived differences or challenges with the new system.

- **No Difficulties Experienced:** Explicit statements that the student did not experience any difficulties.
- **No Perceived Difference:** Perception that the learning experience in the classroom was unchanged.

3.4.2. Teacher Interviews Data Analysis

Teacher interview data from this study was also analyzed through a qualitative content analysis approach, which was selected for its systematic yet flexible means of analyzing textual data through codes developed from the participants' stories and linguistic

choices (i.e., classifying themes). This approach fits well with this research, where this study addressed the perspectives and lived experiences of English language teachers after a significant policy change by the Turkish Ministry of National Education (MoNE). This policy stated that 50% of students' final English grades would now depend on assessment value assigned to speaking and listening.

Qualitative content analysis is especially useful when the intention is to compare patterns of meaning from open-ended responses while acknowledging the meanings arising from participants' deep and contextual viewpoints. This approach does not highlight word counts or pre-defined variables but focuses on the meaning units as they occur in their sociocultural and institutional meanings. This categorizing and interpreting of teacher interview data is essential if researchers want to capture the reflective voice of teachers (i.e. teachers in this study reflect on their identity in their own words, without researcher interference) especially in relation to formative professional identity and workload in light of policy changes, especially the nature of assessment practices that could largely determine a students' final grade, or the limits to their pedagogical agency.

Seven English language teachers from three public middle schools in Çanakkale province were interviewed using semi-structured interviews to elicit narratives that were detailed and story-like and focused on their lived experiences. Each teacher's reflections about the assessment reform were based on their own lived experience, as each teacher had engaged with the revised assessment for at least one semester, which mitigated potential hypotheticals. The interview prompts focused on three broad areas: classroom experiences, assessment challenges for them and their students, and the broader school context.

The appropriateness of qualitative content analysis for this data set is further warranted because qualitative content analysis has inductive logic. Erlingsson and Brysiewicz (2017) note that content analysis captures the potential in which researchers move from raw data to more abstract conceptual themes, while remaining true to the language and meaning used by participants. This analytic pathway was crucial in this study, as teachers often blended affective, descriptive, and critical tones in their responses, ranging from optimism about student engagement to frustration over administrative burdens, and lack of clarity in assessment.

Importantly, the analysis remained open to emergent categories but was guided by existing theoretical insights from the fields of language assessment and teacher cognition. Key literature on washback effects (Cheng et al., 2008), teacher mediation in assessment (Looney, 2011), and policy enactment in language education (Priestley et al., 2015) informed the interpretive stance of the researcher. This theoretically sensitized approach helped ensure that the coding process captured both immediate practical concerns and deeper pedagogical or institutional implications. Another justification for the use of content analysis was its flexibility in dealing with heterogeneous expression styles. Some teachers provided structured, metalinguistic accounts of their experiences, while others spoke in more narrative and emotive terms. A rigid coding or frequency-based approach would have risked flattening these differences. Instead, by adopting an interpretive content analysis framework, the study was able to preserve variability while still identifying patterns and categories across cases.

In summary, the teacher interviews provided a crucial perspective into the lived experiences of English language teachers during a significant assessment reform. Their reflections shared the practical understanding of the process of applying this new system and details about their emotional and pedagogical work and the inherent systemic stressors. The following sections describe the coding process used with the data, the steps in MAXQDA taken to complete this analysis, and finally the justification for developing the themes and codes.

Coding Process

The coding process for data collected through teacher interviews involved an inductive qualitative content analysis method that included careful reading of textual data to find patterns, trends, concepts, interpretative patterns, and significant categories. The coding process aimed to build a transparent and theoretically based coding system that accurately reflected the English language teachers' understanding and reflections of the implementation of the new English assessment regulation introduced by the Ministry of National Education (MoNE).

The initial phase of the analysis began with multiple readings of the interview transcripts. As it allows to achieve immersion in the data, allowing the researcher to grasp the breadth and depth of responses (Lichtman, 2023). Each interview was treated as an individual unit of analysis, and within each interview, individual comments or statements were segmented into meaning units when they presented distinguishable ideas. For example, a teacher's single paragraph might contain a reflection on curriculum misalignment, a concern about speaking anxiety, and a suggestion for increasing class hours all of which were coded under separate subcategories.

As the coding developed, initial codes were generated from teachers' explicit and implicit mentions of their own classroom experience, institutional challenges, emotional concerns, and pedagogical approaches. Codes were generated first with in vivo language (i.e., the teachers' own words) to retain the true meaning (Saldaña, 2021). Codes were then conceptually refined and grouped into broader themes through constant comparative analysis, where the researcher returned to the data to identify their meaning. For instance, a statement such as "Students get stressed and anxious because they don't fully know what criteria they will be evaluated against" was both designed as "Student Anxiety and Stress" and "Unclear Assessment Criteria" since they worked to highlight an emotional focus and transparency with assessment.

In order to improve analytical validity, the researcher triangulated at several levels: (1) provided a cross-case comparison to teachers in different schools, (2) revisited the data iteratively when coding, and (3) participated in peer debriefing with two independent researchers in the field of ELT to assess the clarity and consistency of the code list. This validation process was useful in diminishing researcher bias and ensuring the codes were a reflection of shared patterns rather than singular interpretations (Nowell et al., 2017).

The coding process was inherently hierarchical within MAXQDA, where main themes (parent codes) incorporated more specific subthemes (codes), and finally more specific subcodes. For instance, the theme Challenges and Difficulties Encountered included sub-themes that included Administrative Burden, Time Constraints, Class Size Issues, and Insufficient Teacher Support. The process of developing a hierarchical structure was not linear; new codes had the potential to create new considerations regarding completed

categories, such as combining existing categories or further subdividing them to improve conceptual clarity.

Furthermore, sensitivity to literature informed the development of the several codes. Codes were developed inductively before and throughout the coding, the researcher considered prior work about assessment reform (Looney, 2011), washback (Green, 2007) and teacher agency (Priestley et al., 2015), especially when considering comments that had greater pedagogical implications. For instance, in the quotation where a teacher suggested flexible scheduling of oral exams, it was coded as a request for a procedural change, as well as being connected to the larger themes of teacher agency and institutional flexibility.

In summary, the coding process was based on a balance of empirical rigor and interpretive depth. The aim of the coding was to accurately reflect the lived experiences of teachers yet also develop an analysis framework that would allow for meaningful examination of the thematic interpretation of data. The final code system provides the foundation for the later MAXQDA analysis and visual representation of significant patterns across participants, as presented in the next section.

MAXQDA Analysis Procedure

The audio recordings of interviews were transcribed manually into text documents beforehand and saved as singular documents. These documents were imported using the import documents functionality in MAXQDA. Each teacher's interview was identified with a pseudonymous identity (e.g., Teacher 1, Teacher 2) for participant anonymity, while also allowing comparisons across individuals and institutions. Demographic variables were also added (e.g. years of experience, location of school, grade level being taught) into the "Document Variables" menu for possible future cross-sectional analysis, should it be warranted by patterns based on contextual elements.

The Code System panel consisted of a code tree that contained the 'emerging' themes and their sub-themes hierarchically. Parent codes were developed to capture the higher-order thematic domains such as *Suggestions*, *Perceived Benefits*, and *Challenges and Difficulties Encountered* while child codes captured more specific phenomena within these domains,

such as *Administrative Burden*, *Student Anxiety and Stress*, *Unclear Assessment Criteria*, and *Motivation Increase*. These arrangements allowed for both vertical (thematic depth) and horizontal (comparative range) analysis of the data set.

The code system was developed incrementally throughout the process of analysis. Initially, it was utilized in vivo codes to preserve teacher voices in the data set - for example, *Insufficient Class Hours* and *Student Anxiety and Stress* correspond to in vivo quotes like "I think we need more hours if we are doing the speaking exams properly", and "Students feel anxious because they don't know how they are going to be assessed"; as with each clustering of the quotes, the in vivo codes abstracted to higher-order conceptual categories. The code system was improved by combining similar codes, renaming unclear ones, and keeping the overall structure consistent and easy to follow.

MAXQDA's visualization features were integral to the analytic process, particularly in identifying code frequencies, co-occurrence patterns, and the distribution of themes across cases. Several core tools were employed:

- **Code Matrix Browser**: This function was used to map the presence of each code across individual teacher interviews, allowing the researcher to detect patterns such as which themes were common to all participants (e.g., *Student Stress*) and which were more context-specific (e.g., *Return to Previous System*).

- **Code Frequency Table**: This tool provided a numerical summary of how frequently each code appeared across the dataset. For example, *Insufficient Teacher Support* and *Inconsistency Across Schools* were among the most frequently coded items, indicating systemic-level challenges in policy implementation. It also allowed for the presentation of the data through various visuals such as pie charts or bar charts.

- **Summary Table**: For each main theme, the Summary Table function was used to create concise explanatory summaries that synthesized the key ideas under each code. These summaries were later used to support the thematic analysis section and provided source-linked evidence from teachers' responses.

During the coding stage, several iterations of revisions were made to the structural code as new considerations arose throughout that process. The researchers made extensive

use of MAXQDA's memo feature, to record decisions related to coding, reflections about the boundaries of the themes, or analytic uncertainties. For instance, one memo justified the decision to keep *Coursebook Content Mismat*ch and *Curriculum Misalignment* as separate codes, as they referred to different institutional factors national policy as opposed to decision made by the publisher/design considerations.

Code comments and document memos were also designed for the purpose of keeping track of potential tensions between teachers' individual perceptions and structural constraints. The reflexive layer provided an additional degree of analytical rigor and ultimately bolstered grounded interpretive practices that were sensitive to context.

Rationale for Theme and Code Development

Theme 1: Challenges and Difficulties Encountered

This theme captures the variety of challenges teachers faced in implementing the new English assessment policy. These challenges covered institutional, pedagogical, and affective challenges that indicate both structural and psychological challenges experienced in transitions.

- **Administrative Burden:** Increased demands on administration in the form of scheduling, recording, evaluating, and documenting speaking and listening exams in a limited amount of time and without institutional support.

- **Student Anxiety and Stress:** Stress and emotional pressure placed on students in speaking and listening tasks, often from performance anxiety, lack of experience with an assessed situation, or general nervousness about assessment.

- **Unclear Assessment Criteria:** Ambiguity regarding the standards and expectations by which the speaking and listening performances are evaluated that lead to inconsistencies and confusion.

    - *Insufficient Teacher Support:* Lack of adequate guidance, sample materials, or training on how to apply the new criteria effectively.

- *Inconsistency Across Schools:* Variation in how different schools or teachers apply assessment criteria due to the absence of centralized standards, resulting in perceived unfairness.

- **Curriculum Misalignment:** The disconnect between what the national curriculum asks for from learners and what the new assessment framework represents as acceptable competencies.
    - *Coursebook Content Mismatch:* Incongruity between the textbook materials and the communicative objectives of the speaking and listening components.
    - *Inadequate Coursebook Activities:* Insufficient or ineffective speaking/listening tasks in the coursebooks that fail to prepare students for performance-based exams.
- **Time Constraints:** Limitations caused by restricted instructional time allocated for communicative skill development and assessment.
    - *Insufficient Class Hours:* Lack of sufficient English lesson hours in the weekly schedule to accommodate both curriculum coverage and exam preparation.
    - *Unit Overload:* Excessive number of textbook units to be completed within the term, leaving little time for skill-based teaching or formative speaking practice.
    - *Class Size Problems:* Large class sizes that hinder one-on-one oral assessment, group speaking activities, and effective feedback provision.

Theme 2: Perceived Advantages

This theme consists of teachers' positive reflections on the changes to the assessment policy, considering perceived pedagogical and affective advantages. These codes suggested aspects where teachers felt the new system provided better alignment with communicative goals of teaching, improved attention from students, and increased awareness of language.

- **Conceptually Positive:** The belief that the new system is pedagogically sound and reflects real-life use of language.
    - *Increased Awareness:* Students became more conscious of the practical use of English and recognized the value of communication skills.
    - *Encouraging Speaking Practice:* The reform encouraged teachers and students to engage more in speaking tasks in the classroom.
- **Improved Student Skills:** General perceptions that students' language performance improved by the new evaluation model.
    - *Listening Skill:* Improvement in students' ability to comprehend spoken English through regular exposure and exam preparation.
    - *Speaking Skill:* Observable development in students' fluency, vocabulary range, and willingness to speak in English.
- **Motivation Increase:** Students appeared to be more enthusiastic and had an improved disposition toward English lessons from the new assessment focus.

Theme 3: Suggestions

This theme consists of a range of teacher-generated ideas and reflective aspects of improving the implementation of the revised English assessment system. The ideas reflect a need for clearer structures, more efficient institutional planning, and better curriculum design responsiveness, so teachers and students can be better supported.

- **Proposed Solutions:** Broad suggestions provided by teachers to address the operational and pedagogical challenges of the new system. These include systemic, curriculum-based, and school-level reforms.

- *Informing Teachers*: Emphasis on the need for comprehensive professional development, clear communication, and training sessions prior to implementing the new assessment model.
- *Prior Research*: Teachers suggested that before large-scale changes are introduced, pilot studies or needs analyses should be conducted to identify potential issues.
- *Clear Assessment Criteria*: The demand for standardized rubrics and explicit evaluation descriptors to ensure fair, consistent, and transparent grading.

- **Curriculum Revision:** Recommendations to align the curriculum with the new exam components and reduce pressure on classroom time and content coverage.
  - *Return to Previous System:* A few teachers expressed a desire to revert to the former system due to the difficulties encountered.
  - *Updating Coursebook:* Suggestions to revise the current coursebooks to better support speaking and listening development.
  - *Reducing Unit Load:* Calls to decrease the number of textbook units to allow for deeper engagement with language skills.
- **Increasing Class Hours:** Proposals to expand weekly English class hours to accommodate communicative skill development, exam practice, and feedback sessions.

Theme 4: Individual Solutions

This theme represents teachers' localized and personally cultivated strategies for navigating the challenges provided by the new assessment framework, and represents agency and creativity in response to time, resources and institutional uncertainty.

- **Flexible Scheduling:** Determining assessment times based on the conditions of the school or students, such as spreading exams over a few days or incorporating assessment within normal lessons.

- **Additional Support for Students:** Offering students extra support outside of or during class time, such as feedback, practice or personal encouragement.

- **Individualized Assessment:** Modifying assessment style or expectations according to students' proficiency levels or emotional state to allow for fairer and more supportive assessment.

3.4.3. Intercoder Reliability

In order to uphold the reliability and consistency of the qualitative data analysis, an intercoder agreement process was performed using MAXQDA 2020 software. As Landis and Koch (1977) emphasized, the measurement of agreement between independent raters is essential for assessing the reliability of categorical coding beyond chance effects. Assessing intercoder reliability on a subset of the dataset is common practice in qualitative content analysis, provided that the subset adequately represents the range of themes and categories identified in the data (Cheung and Tai, 2021). In line with this view, two coders independently analyzed approximately 30% of the entire student dataset ($n = 89$), and %40 of teacher dataset ($n = 3$) using the same coding framework for both groups accordingly. The dataset waere purposely selected to be a representative segment of the entire data, by including examples from all main themes, while also providing proportional coverage of the entire data set.

The two coders' comparison resulted in 220 agreements out of 264 coding decisions for an overall agreement rate of *83.33%*. The calculated coefficient of *Cohen's Kappa (κ)* was 0.81, which indicates substantial agreement based on the values given by Landis and Koch (1977). This level of agreement indicates that the coders clearly understood the code system and worked within the same interpretive framework.

Some codes showed small discrepancies in agreement across a few of the codes, primarily in subjective descriptors such as *Better Learning* and *Feeling of Improvement*. In general, for most of the codes (e.g., *Improved Grades*, *Confidence in Speaking*, and *Listening Skills Improvement*), the agreement was at *90%* or above. Thus, it was concluded with a high level of certainty that there was a high level of agreement among the coders across triangulation and coding reliability.

3.4.4. Trustworthiness of the Study

To ensure the trustworthiness of the qualitative findings, several strategies were employed throughout the research process. In qualitative research, credibility is closely linked to the accuracy and authenticity of the participants' experiences as represented in the data (Lincoln and Guba, 1985). In this study, credibility was established through prolonged engagement with the data, triangulation between teacher interviews and student survey responses, and the inclusion of direct quotations to illustrate the participants' voices. The coding process in MAXQDA was conducted meticulously, and data were revisited multiple times to verify emerging patterns. The researcher's continuous reflection on personal assumptions and biases further enhanced the credibility of the interpretation.

Transferability was achieved by providing a detailed description of the research context, including information about the participating schools, the number of teachers and students, and the educational setting in Çanakkale. Such detailed contextualization allows readers to determine whether the findings can be applied to similar educational contexts (Shenton, 2004). The inclusion of representative excerpts from participants' responses also helps convey the diversity of experiences and ensures that the findings are not overly generalized beyond their original context.

Dependability was ensured by maintaining a clear and consistent research procedure throughout the study. All stages, including data collection, transcription, translation, and analysis, were documented systematically to establish an audit trail. Using MAXQDA for coding and thematic analysis provided an organized and transparent data management process. Additionally, to ensure the consistency of coding, themes were reviewed and refined

several times to confirm their alignment with the original data. This approach supports the replicability of the study's procedures, even if not its exact results.

Confirmability was strengthened through the researcher's effort to minimize bias and to allow the findings to emerge from the data rather than from personal interpretations. Direct quotations from teachers and students were presented to show how themes were grounded in the participants' actual words. Furthermore, all analytical decisions, including theme development and code refinement, were recorded to establish transparency in the analytic process. By maintaining neutrality and grounding interpretations in evidence, the study ensured that the findings reflected the participants' realities rather than the researcher's expectations.

3.5. Chapter Summary

In this chapter, the research design, context, participants, data collection procedures, and also data analysis procedures applied in the current study have been described. A qualitative research design was adopted as the study aimed to investigate reflections, perceived advantages and disadvantages, and difficulties of seventh-grade students and English language teachers on the recently introduced English assessment regulation in Türkiye.

The study took place at three public middle schools in Çanakkale, and data was collected through open-ended student surveys from 89 students and semi-structured interviews conducted with 7 English teachers. Primarily, 92 students volunteered to participate in surveys, three student survey data were excluded as a result of answers being too vague and not holding statistical value. Ethical procedures were adhered prior to data collection, including receiving consent from the Turkish Ministry of National Education and consent from parents for student participation. Development of both instruments was underpinned by literature and expert consultation. The student surveys and teacher interviews were all conducted in Turkish to accommodate students' comfort and ease of use.

Collected data was analyzed through qualitative content analysis using MAXQDA, following a structured coding and thematization process. The chapter also described the

rationale for theme and code development based on both inductive interpretation and thematic relevance. The next chapter presents the findings of the study, organized according to the three research questions and supported by frequency-based visualizations and illustrative participant quotes.

CHAPTER FOUR
RESULTS

4.1. Introduction

This chapter provides the study results based on the findings from the thematic analysis of qualitative data collected from student surveys and teacher interviews. This chapter aims to systematically present and document the results in direct response to the research questions identified in Chapter I:

1. *What are the general reflections and perceptions of teachers and students regarding the new regulation?*
2. *According to teachers and students, what are the perceived benefits and drawbacks of the new regulation?*
3. *What specific challenges and difficulties have been experienced in the implementation of the new regulation?*

A total of 89 valid student survey responses from 7th-grade students were analyzed. Even though the survey was distributed to 92 students, three responses were excluded since they held no statistical value. In addition, semi-structured interviews were conducted with 7 English language teachers to take a deeper look at their experiences with the new assessment system. Together the surveys and the teacher interviews allow for a robust understanding of both the students and teachers experience within the new regulation.

The data analysis phase used MAXQDA, which allowed for systematic coding and thematic structures to be created. The thematic analysis was based on an inductive approach where codes and themes were first primarily derived from the data, along with some supportive literature and assistance from experts in the field. Therefore, from the student data survey, five different themes emerged (*Skills Development*, *Academic/Personal Progress*, *Negative Attitudes/Experiences*, *Challenges/Difficulties*, and *Neutral/No Perceived Change*). The teacher interview data produced parallel themes in the context of the teacher perspectives which consisted of four main themes (*Perceived advantages*, *Challenges and Difficulties Encountered*, *Individual Solutions*, *Suggestions*). Each theme

was divided into subcodes to capture more nuanced responses; frequency tables were then created to illustrate and summarize the codes within participants.

In this section, general overview of the themes of both student and teacher data are presented using two tables (Table 1 and Table 2). These tables illustrate the main thematic structure and their frequencies before the research question sections in which detailed analysis are supported by figures such as pie charts and bar charts.

Table 1

Student survey themes

Theme	**Students (n)**	**Mentions (f)**
Challenges/Difficulties	58	84
Academic/Personal Progress	51	66
Skills Development	49	74
Negative Attitudes/Experiences	41	68
Neutral/No Perceived Change	30	31

Table 2

Teacher interview themes

Theme	**Teachers (n)**	**Mentions (f)**
Perceived Advantages	7	18
Challenges and Difficulties Encountered	7	76
Individual Solutions	4	18
Suggestions	3	12

The chapter is divided into three sections according to the research questions. RQ1 provides a summary of key reflections and views from students and teachers and addresses the impressions of the shift to a communicative and skills-based approach to assessment.

RQ2 reviewed the positives and negatives of the system, and RQ3 examined the specific difficulties the teachers experienced implementing it. The combination of quantitative code prevalence with qualitative excerpts offers both extents and intensities of the participants' experiences during the shift.

4.2. Research Question 1

4.2.1. Student Surveys

Research question one (RQ1) was concerned with how the students and teachers understood and evaluated the new English assessment regulation. The current section will concentrate on the students' responses, which represents a key account of how the regulation was experienced by the students in the context of their English language learning process. A total of 89 students indicated that they had had a new assessment regulation for their assessment and provided us with their thoughts based on open-ended survey response items. The answers were analyzed with thematic coding in MAXQDA, and we were able to identify five full-scope themes from studying students' survey responses to these open-ended qualitative responses: *Academic/Personal Progress*, *Skills Development*, *Challenges/Difficulties*, *Negative Attitudes/Experiences*, and *Neutral/No Perceived Change*. These themes were derived inductively from students' unstructured qualitative survey responses to capture as many students' emotional experiences and perceptions as possible.

As exemplified in Table 1, patterns emerge from the distribution of mentions (*f*) and the number of students (*n*) associated with each category in the data. The theme *Problems/Difficulties* stood out clearly ($n = 58$; $f = 84$), suggesting that a large number of students faced challenges in adjusting to the new assessment system in their English classes.

After that, many students mentioned *Skills Development* ($n = 49$; $f = 74$) and *Academic/Personal Progress* ($n = 51$; $f = 66$), mostly in a positive way. Their answers showed that the new policy helped them improve their language skills or feel more motivated. On the other hand, the theme *Negative Attitudes/Experiences* ($n = 41$; $f = 68$) pointed to emotional struggles like feeling anxious, discouraged, or unsure.

Lastly, *Neutral/No Perceived Change* (n = 30; f = 31) captured students' responses with the expression that their learning changed in no significant way in neither positive nor negative categories.

The following sections will provide details for each theme, including constituent codes and subcodes, and examples of exempted student responses that illustrated the principal findings.

Academic/Personal Progress

The *Academic/Personal Progress* theme depicts students' awareness of progress in their English learning performance and personal development as a result of the altered assessment regulation. There were 66 references made by 51 of the students, which means that more than 50% of the participants identified positive impacts in the form of improvement in learning, grades, or study habits. Figure 2 shows that the predominant code was *Improvement in Learning* (n = 39), followed by *Improved Grades* (n = 19), *Improved Self-Study Habits* (n = 4), and *Increased Motivation* (n = 1).

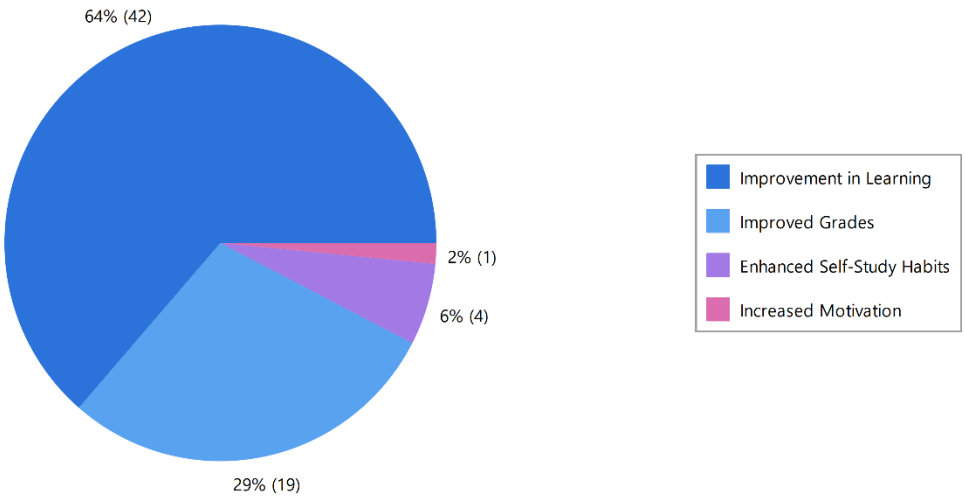

Figure 2. Student perception of academic and personal progress experienced

Student Quotes:

- **S5:** “I have learned more English compared to the past; the speaking and listening tasks motivated me to study harder.”
- **S12:** “My English grades have increased significantly because I started preparing more seriously for the listening and speaking parts.”
- **S34:** “I feel like I understand the lessons better now; I have started to practice more at home.”
- **S51:** “The new exams made me study by myself and focus more on improving my vocabulary.”

The high prevalence of *Improvement in Learning* suggests that the majority of students found the communicative assessment tasks useful for active engagement and practice. The references to *Improved Grades* suggest that motivation, regarding the assessment, was also important to their overall performance. Similarly, *Enhanced Self-Study Habits* and *Increased Motivation* were identified less frequently but nevertheless reflect further indicators of shifting students' autonomous learning behaviors, and the degree to which students are being prompted to be responsible for their own learning in this system.

Skills Development

The *Skills Development* theme illustrates how the new assessment regulation allowed students to develop their communicative competence, particularly with regard to speaking and listening. This theme created a noticeable impression on the participants of the study as it was found 74 times with 49 students' answers (n = 49; f = 74). This indicates that nearly half of the participants noticed improvement in their speaking and listening skills. The most common code in the *Skills Development* theme was *Speaking Skill Improvement* (n = 48;), followed by *Listening Skill Improvement* (n = 20), and *Memorization and Retention* (n = 5) as shown in Figure 3 below.

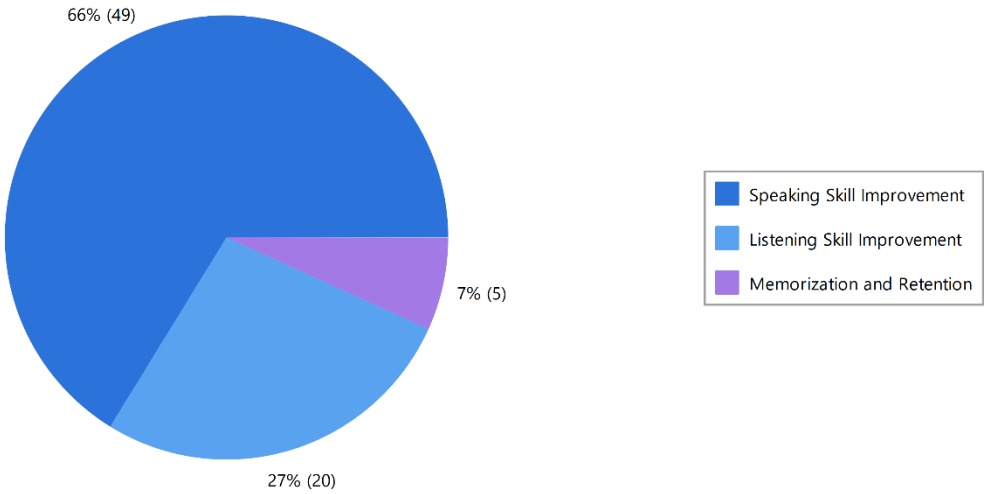

Figure 3. Students' perceived skill development

Student Quotes:

- **S8:** "I have improved my speaking because we practiced speaking in class more frequently than before."
- **S21:** "Listening exercises helped me understand English conversations better."
- **S37:** "Since the new exam includes speaking, I started memorizing words and phrases more effectively."
- **S45:** "I feel more confident when talking to others in English because of the speaking tests."

The evident majority of *Speaking Skill Improvement* indicates that the speaking component of the new assessment has benefits for students with oral proficiency and confidence directly, while *Listening Skill Improvement* highlights the benefit of authentic listening exercises as students perceived to be good for comprehension and interaction. *Memorization and Retention* code demonstrates that students used strategies in memorizing important phrases or vocabulary for assessment performed well which indirectly contributed to developing their language skills.

Challenges and Difficulties Encountered

Overall, the *Challenges/Difficulties* theme represents all the challenges and difficulties students experienced while enacting the new assessment regulation. Given that it had 84 mentions across 58 students, it was the most mentioned theme. As indicated in Figure 4, the broad challenges included *Assessment-Related Challenges* (*n* = 24), *Difficulty with Listening Tasks* (*n* = 22), *Difficulty with Speaking Tasks* (*n* = 21), and *Difficulty in Adaptation* (*n* = 10).

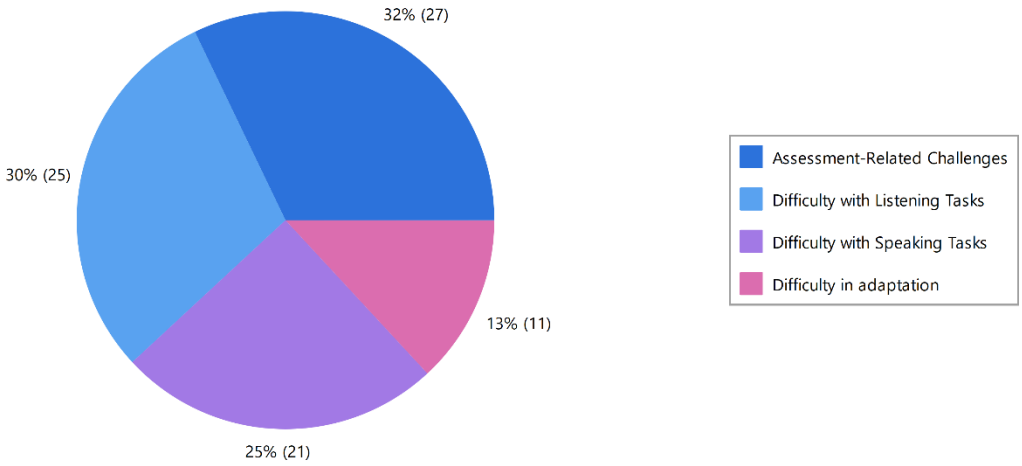

Figure 4**.** Challenges and difficulties faced by students

Student Quotes:

- **S16:** “The listening part was quite difficult for me because I couldn’t understand the recordings well.”
- **S25:** “Speaking in front of the class made me nervous, and I couldn’t express myself properly.”
- **S40:** “I found it hard to prepare for both listening and speaking exams at the same time.”
- **S57:** “It was difficult to adapt to this new system because it was different from what we were used to.”

The results for *Assessment-Related Challenges* indicate that students faced challenges related to the structure of the new exams, struggling with multiple assessment components, and preparing for communicative tasks. *Difficulty with Listening Tasks* and *Difficulty with Speaking Tasks* are primarily skill-based and involved, in particular, anxiety, pronunciation, and comprehension issues. *Difficulty in Adaptation* suggests that students experienced a considerable adjustment process from a traditional exam-oriented format to a communicative assessment model.

Negative Attitudes / Experiences

The *Negative Attitudes/Experiences* theme captures the students' negative perceptions and emotional experiences of the new assessment regulation. It was mentioned 68 times by 41 students, which emphasizes that a sizable group of learners perceived the new system as stress, anxiety, or a perceived or real lack of progress. Figure 5 shows that the code with the most mentions is *Anxiety and Stress* (*n = 27*), *No Improvement Felt* (*n = 17*), *Decreased Grades* (*n = 14*), *Loss of Motivation* (*n = 5*), and *Inhibited Learning* (*n = 2*).

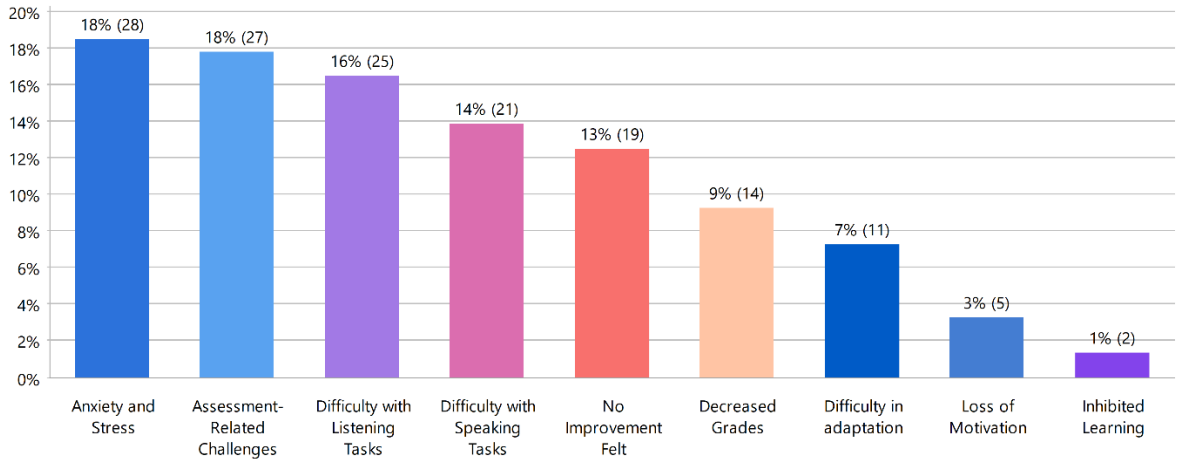

Figure 5. Negative attitudes and experiences of students

Student Quotes:

- **S14:** “I get nervous when I have to speak in front of the class; it makes me forget what I want to say.”
- **S28:** “I don’t think my English has improved with this new exam system.”
- **S42:** “My grades went down because I’m not good at listening exams.”
- **S56:** “The new system made me feel less motivated since I am always stressed about speaking.”

The prevalence of *Anxiety and Stress* indicates that the speaking and listening sections are notable performance stressors for many students. The mentions of *No Improvement Felt* and *Decreased Grades* seem to point to a proportion of learners seeing the new system as an impediment rather than a vehicle for improvement. More infrequent, but still notable, are the mentions of *Loss of Motivation* and *Inhibited Learning*, which can indicate how adverse emotional experiences may restrict student voluntary engagement in language task completion.

Neutral/No Perceived Change

The *Neutral/No Perceived Change* theme outlines the responses of students who stated that the new assessment regulation had a neutral or no effect, either positively or negatively, on their experience of learning English. This theme was mentioned 31 times in the responses of 30 students, showing that a small but significant number of learners perceive the new assessment regulation as having little or no impact. Figure 6 shows the codes under this theme which are *No Perceived Difference* (n = 19) and *No Difficulties Experienced* (n = 12).

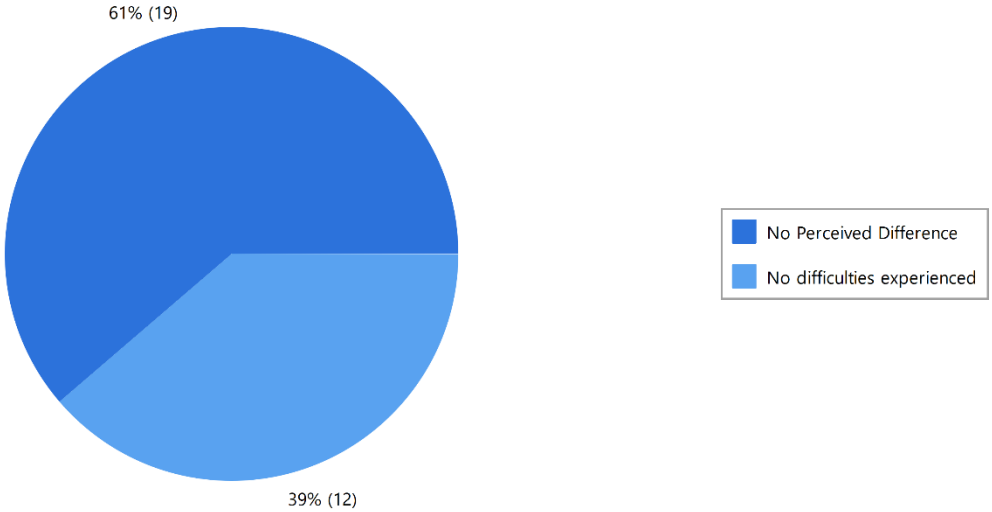

Figure 6. No difference or difficulty experienced from student perspective

Student Quotes:

- **S3:** “It did not help anything at all.”
- **S17:** “I haven’t faced any difficulties with the new exams.”
- **S33:** “It feels almost the same as before, I didn’t notice much change.”
- **S60:** “I don’t think the new exams made any difference in my learning.”

The reference to *No Perceived Difference* reflects that, although a change occurred regarding the assessment, there were instances where students believed that they felt there was no difference in their learning, in classroom practices, and did not have any difficulties under the new regulation. The reference to *No Difficulties Experienced* indicates a similar idea; some students stated they did not find the new assessment regulation greatly different from what they were doing previously and therefore did not experience any challenges in particular. Overall, this suggests a neutral stance since the students did not see either benefits or disadvantages to the assessment experience.

4.2.2. Teacher Interviews

The first research question examined what English teachers thought and felt about the new assessment. The interviews were analyzed in MAXQDA, and four main themes emerged: *Perceived Advantages*, *Challenges and Difficulties Encountered*, *Individual Solutions*, and *Suggestions*. These themes were built directly from teacher responses. Subcodes under each theme showed common concerns or ideas that repeated across interviews.

As shown in Table 2, the number of teachers (*n*) and total mentions (*f*) varied across themes. The most frequent one was *Challenges and Difficulties Encountered* ($n = 7$; $f = 76$), showing that most teachers talked about the problems they faced. *Suggestions* was the least mentioned theme ($n = 3$; $f = 12$) yet still offered useful ideas about how to improve the implementation of the regulation theoretically and practically. Both *Perceived Advantages* and *Individual Solutions* were mentioned the same number of times ($f = 18$), which suggests teachers noticed both helpful aspects and came up with ways to manage problems. The following parts of this section examine each theme in more depth, using interview excerpts to illustrate what the teachers said.

Perceived Advantages

The theme *Perceived Advantages* shows that teachers identified several positive aspects of the new regulation. All seven teachers pointed out at least one perceived benefit. The most common ideas were *Conceptually Positive* ($n = 4$) and *Improved Student Skills* ($n = 3$). These were followed by *Encouraging Speaking Practice* ($n = 2$), as shown in Figure 7.

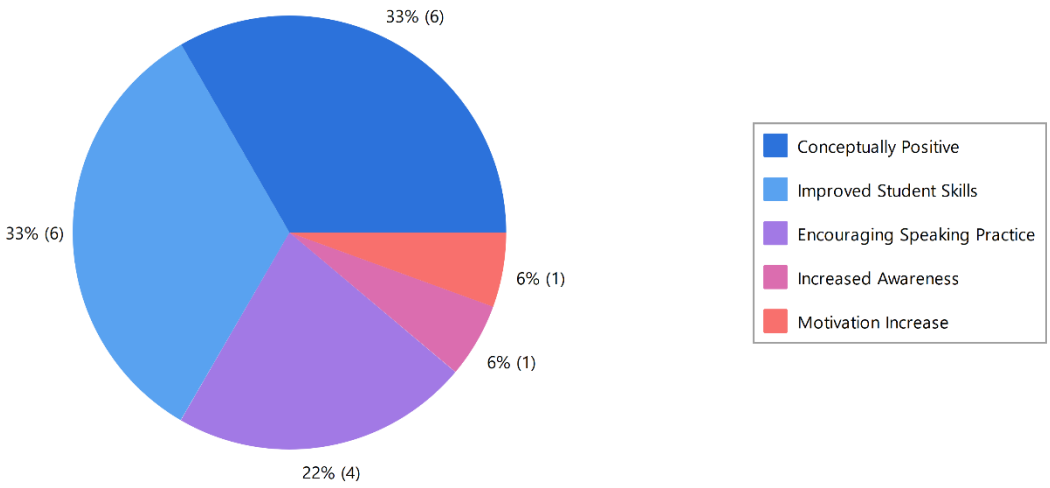

Figure 7. Teachers' perception of advantages

Teacher Quotes:

- **T2:** *"This system is conceptually strong because it aligns with how English is actually used in daily life."*
- **T5:** *"There is visible improvement in their ability to express ideas verbally."*
- **T3:** *"They now know that speaking will affect their grades, so they take the activities more seriously."*
- **T7:** *"They now understand that English is not only about writing or grammar but also about being able to communicate effectively."*

These results suggest that many teachers agreed with the main idea of the change. Several participants believed that the policy fit well with the goal of helping students use English in real life. Some also said that speaking exams helped students get more involved in class and speak more freely. One teacher noted that students felt more active because they were now being tested on their speaking, not just grammar or vocabulary. This shows that the regulation encouraged more practice and real communication in class.

Challenges and Difficulties Encountered

The theme Challenges and Difficulties Encountered points to the real problems teachers faced with the new system. All seven teachers talked about at least one difficulty. Some issues came up more than once. Figure 8 demonstrates how often each problem was mentioned and by how many teachers.

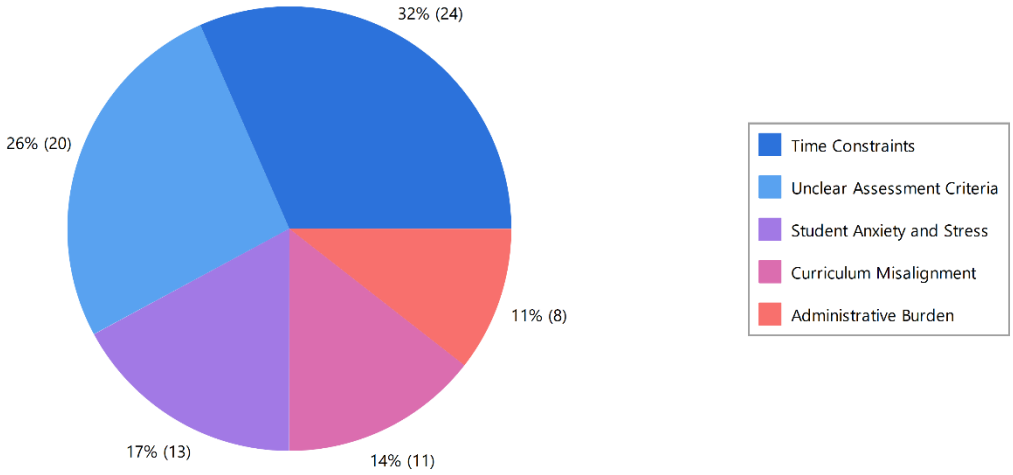

Figure 8. Challenges and difficulties encountered by teachers

Teacher Quotes:

- **T2:** *"Students often get anxious during speaking and listening tasks, which affects their performance significantly."*
- **T3:** *"The assessment criteria are not always transparent, and this creates confusion both for teachers and for students."*
- **T4:** *"The coursebook content does not fully align with skill-based tasks, which makes exam preparation challenging."*
- **T6:** *"We cannot find enough time to balance the curriculum, test preparation, and interactive activities due to limited class hours and large units."*

The most common problem was *Time Constraints* ($n = 6$). Teachers said they do not have sufficient time to fit in the speaking and listening exams along with the rest of the

curriculum. *Unclear Assessment Criteria* ($n = 5$) was also a principal concern. Many teachers mentioned they didn't know how to grade the speaking part properly because there was not a clear rubric. Another common issue was *Student Anxiety and Stress* ($n = 6$). Teachers noticed that the students were often nervous during the oral exams, which influenced their performance in a negative way.

Teachers also talked about *Administrative Burden* ($n = 4$), mentioning the extra paperwork and management of the examinations considering the inconveniences that occur made things difficult. Finally, *Curriculum Misalignment* ($n = 4$) identified that the textbook content was not well aligned with skill-based tasks and thus the teachers needed to amend the curriculum. Collectively, these issues shed light on how applying the new regulation as a reality was highly challenging.

Individual Solutions

The theme Individual Solutions shows how teachers tried to deal with the new system on their own. This theme was mentioned 18 times by four teachers. Figure 9 presents the main ideas under this theme and how often each code came up.

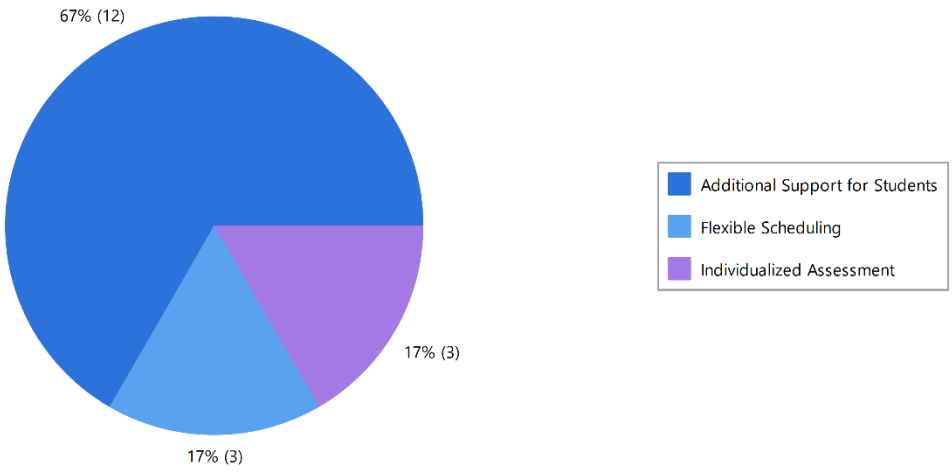

Figure 9. Individual solutions implemented by the teachers

Teacher Quotes:

- **T2:** “For students with high anxiety, I let them speak to me privately, without the presence of their classmates, to help them feel more comfortable during the speaking exam.”

- **T3:** “We share the topics with students beforehand and let them practice in smaller groups or individually to reduce their fear of making mistakes.”

- **T7:** “I try to skip unnecessary parts of the textbook and bring in interactive activities to keep students engaged and manage time more efficiently.”

The most frequent code was *Additional Support for Students* (n = 4). The teachers provided students with additional support before or during the tests and additional speaking practice; or presented relaxing tips to consider before the test or allowed students to do the test alone or in groups. The code *Flexible Scheduling* (n = 2) shows that some teachers found ways to help students who missed the exams or adjust the exam time to avoid student stress. *Individualized Assessment* (n = 2) reflects how teachers changed the way they asked questions based on each student’s level. These frequent mentions show that teachers tried to build a safe space where students could feel less nervous about the exams. *Flexible Scheduling* (n = 2) illustrates how teachers worked around the time limits specifically in big classes. The code *Individualized Assessment* (n = 2) reflects teacher efforts to adjust speaking and listening tasks based on what each student could handle. These examples point to how teachers found their own ways to make the system work more fairly and help students feel included.

The theme *Suggestions* collected the ideas teachers shared to help improve how the new implementation of the regulation might be improved. The theme came up 12 times and was mentioned by 3 teachers. Figure 10 shows the most common codes: *Curriculum Revision* (n = 5), *Clear Assessment Criteria* (n = 1), *Informing Teachers* (n = 1), and *Increasing Class Hours* (n = 2).

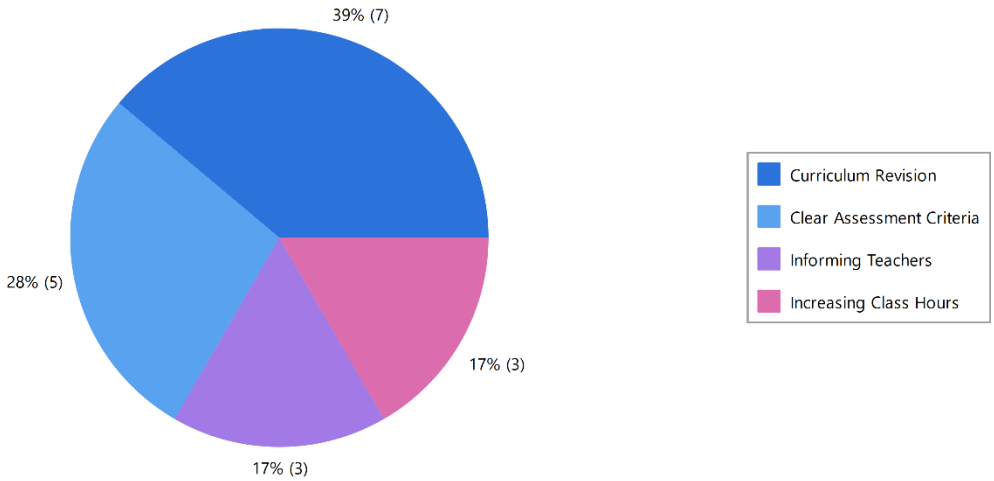

Figure 10. Suggestions proposed by the teachers

Teacher Quotes:

- **T1:** "There should be common criteria for all parts of the exam, including listening, speaking, reading, and writing, instead of leaving everything to the discretion of each school or teacher."
- **T1:** "A possible solution would have been to consult experts and gather ideas from the field before making such changes. I wish they had asked us how this system would work in practice."

The strong focus on *Curriculum Revision* suggests that many teachers want a lighter and more practical curriculum that fits well with the changed and communication-based assessment and evaluation system that was brought with the regulation. The code *Clear Assessment Criteria* points to teachers' need for more stable grading rules which are similar and fair for the varying schools within the educational system. *Informing Teachers* and *Increasing Class Hours* show that better training, informing the teachers prior to the regulation or the reforms to be implemented, and more time is needed. Overall, these suggestions show that teachers are not just pointing out problems, they are also offering real solutions for the optimal benefit of the students and the teachers.

4.3. Research Question 2

4.3.1. Student Surveys

Research Question 2 aimed to identify the advantages and disadvantages of the new English assessment regulation from the students' perspective. Thematic analysis was conducted on the 89 student responses and both positive and negative perspectives were identified. Students identified several positive aspects, mostly related to *Academic/Personal Progress* and *Skills Development*. These included language improvement, more effective study habits, and greater confidence in speaking and listening tasks. The students' expressed negativity was linked to the themes of *Challenges/Difficulties* and *Negative Attitudes/Experiences*, which focused on the students' difficulties regarding the assessment and assessment framework, anxiety and stress while completing performance tasks and the students' comments on what appeared to be a decline in motivation or grades.

As presented in Figure 11, the codes under the benefits dimension, such as *Improvement in Learning* ($n = 39$) and *Speaking Skill Improvement* ($n = 48$), were frequently referenced.

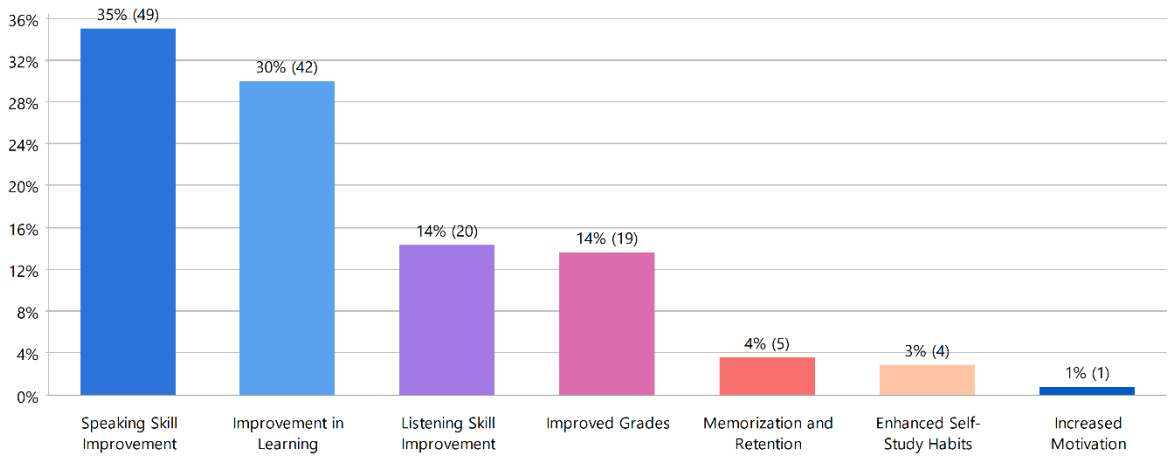

Figure 11. Benefits perceived by the students

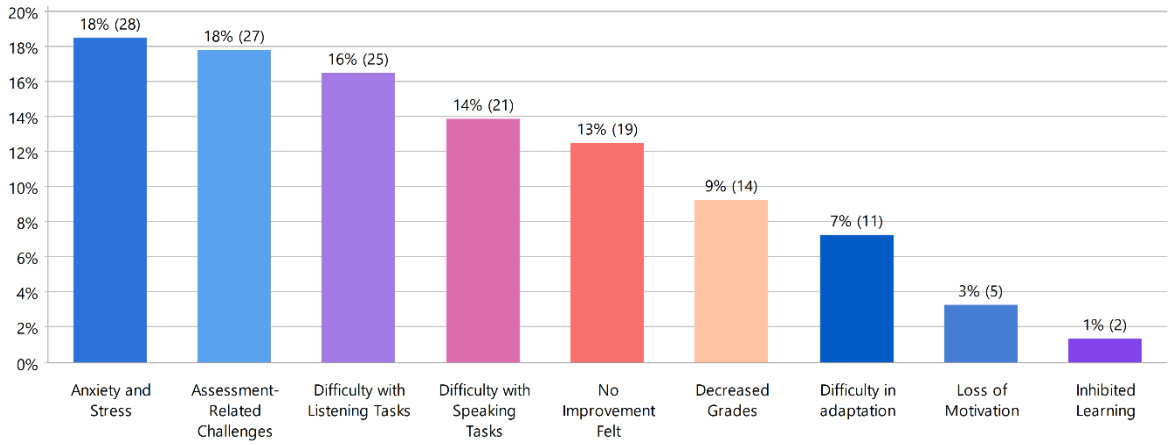

Figure 12. Challenges and limitations students faced

Limitations were particularly notable in codes such as *Assessment-Related Challenges* ($n = 24$) and *Anxiety and Stress* ($n = 27$), as seen in Figure 12. This indicates that while the new regulation allowed many students to develop their communicative skills, there were significant challenges that remain to be overcome before subsequent applications of the new regulation.

Students' Perspective on the Benefits of the Regulation

The benefits identified by students were mostly related to improvements in communication skills and personal learning. As displayed in Figure 11, the most discussed codes were *Speaking Skill Improvement* ($n = 48$) and *Improvement in Learning* ($n = 39$). These results suggest that the speaking and listening components increased interaction and more meaningful language use in classrooms.

Student Quotes:

- **S1:** "Affected in a good way but only because it was useful."
- **S3:** "It was kind of good."
- **S9:** "Influenced my learning in a good way."

- **S9:** “Improved my speaking and listening.”
- **S10:** “It improved my desire to learn English.”
- **S10:** “Mostly improved my memorization of new vocabulary.”
- **S11:** “New evaluation system had a good influence on my English.”
- **S11:** “It helped me learn how to speak English better
- **S14:** “It had a good effect on me. The new exam system made me learn better.”
- **S16:** “It influenced my learning in a good way. I have more interest now.”

The high frequency of *Speaking Skill Improvement* suggests that gains in speaking ability helped increase students’ confidence and willingness to use English in real communication. Also, *Improvement in Learning* and *Listening Skill Improvement* show that communicative tasks had a positive effect on both comprehension and use of the language. There are some less frequently coded data but significant codes such *as Enhanced Self-Study Habits* and *Increased Motivation* that represent the emergence of autonomous learning behaviors.

While some students made generalized and vague comments about improvements from the regulation (e.g., "it was good" or "it helped me"), their comments were much more specific about the improvements in speaking and listening skills. This implies that learners had a greater awareness of their progress for specific skills pertaining to speaking and listening, than to their progress in learning English overall.

Student View of the Limitations of the New Regulation

With *Anxiety and Stress* having so much dominance, it suggests performance-based tasks can invoke significant test anxiety. *Assessment-Related Challenges* reflect issues with preparing for a number of skill-based parts. The standalone code *Difficulty in Listening and Speaking* indicated difficulties for students with their specific skills. The codes *No Improvement Felt* and *Decreased Grades* indicated not all students feel the new system is an

improvement, highlighting the key need to incorporate more balance and support with assessments.

Student Quotes:

- **S12:** “It had a bad influence. It is very difficult to do speaking exams.”
- **S61:** “Its influence was not good; lessons are way more difficult now.”
- **S64:** “With the new system I started to struggle and feel stressed.”
- **S73:** “The new system is effective, but I think it puts pressure on us.”
- **S77:** “It became more difficult in terms of grading.”
- **S19:** “It did not help anything and only stresses me.”
- **S71:** “It made me study harder because if I don’t answer questions in the speaking part, I lose a lot of points.”

While there were several positive outcomes, students indicated that they had a few negative outcomes as a result of the new policy. The most frequently coded were *Anxiety and Stress* (*n* = 27), *Assessment-Related Challenges* (*n* = 24), and *Difficulty with Listening Tasks* (*n* = 22) as indicated in Figure 12. The findings demonstrate that while the system appears to promote communicative skills, it also represents extra stress and challenges for a number of students.

4.3.2. Teacher Interviews

Teachers’ Perspective on the Benefits of the Regulation

The *Perceived Advantages* theme represents the teachers' characterization of the benefits of the new assessment regulation, most importantly the potential it provided in developing the students' communicative capacity and engagement. While the majority of the teachers believed the regulation was conceptually beneficial, many believed the perceived benefits were minor due to contextual constraints such as limited hours coupled with large

class sizes. The most considered codes, as was summarized in Table 9, were *Conceptually Positive* (*n* = 4) and *Improved Student Skills* (*n* = 3), while *Encouraging Speaking Practice* (*n* = 2) was mentioned as meaningful but represented partial benefit.

Teacher Quotes:

- **T1:** “Conceptually it could be good, but in practice, nothing new has been done so far.”
- **T1:** “Yes, it could be beneficial, and I believe it should exist, but not in this way.”
- **T3:** “The advantage is that even a student who has never spoken before can discover themselves during the speaking exam.”
- **T5:** “The speaking component encourages even shy students to express themselves more in class.”

Though these comments emphasize the conceptual benefit of the regulation, teachers were more consistent in pointing out the limitations compared to benefits. Teacher 1's comments illustrate that while the regulation is of conceptual value or positive from a conceptual perspective, in practicality it is not producing the desired outcomes due to the issues that have come with the implementation.

Across all teachers the benefits illustrated were relatively limited in frequency compared to the challenges discussed. For example, *Conceptually Positive* (*n* = 4) and *Improved Student Skills* (*n* = 3) emerged far less than the larger challenge codes which dominated the discussion like *Time Constraints* (*n* = 6) or *Unclear Assessment Criteria* (*n* = 5) suggesting that the teachers acknowledged the regulation's positive aspects, placing greater emphasis on the challenges at the same time.

Teacher View Limitations of the New Regulation

The *Challenges and Difficulties Encountered* theme encapsulates the principal limitations of the new assessment regulation delineated by teachers. The limitations primarily included time management, student stress, uncertainty of criteria, and

inconsistency in the curriculum. As highlighted in Figure 8, the most common challenge was *Time Constraints* (*n* = 6), followed by *Unclear Assessment Criteria* (*n* = 5) and *Student Anxiety and Stress* (*n* = 6).

Teacher Quotes:

- **T1:** "We already have very limited hours, and now we need extra time for speaking exams, which disrupts the schedule."
- **T2:** "Students feel more nervous during these exams, especially when they have to speak in front of everyone."
- **T3:** "There is no clear rubric for evaluating speaking; each teacher ends up grading differently."
- **T4:** "Recording scores, preparing separate exams, and handling paperwork is taking too much time from teaching."

T5: "The curriculum is not designed for communicative assessments; we often have to improvise."

While the teachers have observed some positives of the regulation, the implementation appears to be hindered by restrictions regarding time, uncertainty around grading standards, and inconsistency with the current curriculum. The noted student anxiety and increased workload for teachers strongly suggests that the regulations may need considerable revising in order to be more equitable and less burdensome for teachers and students.

The challenges and limitations were mentioned in far greater quantities than the benefits, and by nearly every teacher. The number of perceived advantages discussed were lesser (e.g., *Conceptually Positive*, n = 4), compared to the theme of Challenges and Difficulties Encountered and its associated barriers like *Time Constraints* (n = 6), and *Unclear Assessment Criteria* (*n* = 5). The themes and codes related to challenges were predominant themes that emerged in the interviews. This suggests the teachers discussed, for the most part, the difficulties and challenges involved in the introduction of the new regulation and its implementation rather than its advantages.

The challenges and limitations were mentioned in far greater quantities than the benefits, and by nearly every teacher. The perceived advantages were fairly limited (e.g., *Conceptually Positive*, $n = 4$), whereas the theme of *Challenges and Difficulties Encountered*, which included barriers such as *Time Constraints* ($n = 6$) and *Unclear Assessment Criteria* ($n = 5$), were major themes in the interviews. This is indicative of the fact that the teachers were predominantly discussing the impediments and logistical challenges associated with the implementation of the new regulation rather than advantages.

4.4. Research Question 3

4.4.1. Student Surveys

The theme *Challenges/Difficulties* was the most common theme ($n = 58$) and included challenges that students experienced including assessment anxiety, problems with listening comprehension and problems with speaking tasks.

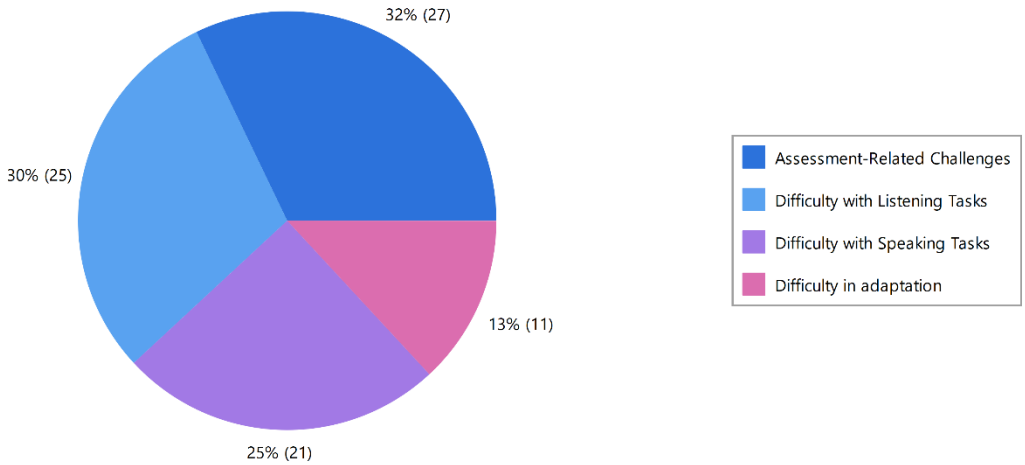

Figure 13. Challenges and difficulties experienced by students

Student Quotes:

- **S2:** "Listening was difficult for me, others were easy."
- **S5:** "It would be better if there were no speaking and listening tests."

- **S6:** “I am an insecure person and that’s why speaking was hard for me.”
- **S8:** “Listening tracks were not audible on smartboard.”
- **S11:** “Conversations in listening tracks are too fast.”

S15: “Grading system was difficult for me because sometimes it felt unfair.”

The most frequently cited codes, such as *Assessment-Related Challenges*, as well as *Difficulty with Listening Tasks*, indicate that students struggled with the exam structure and with specific skill-based tasks. Speed of audio and the conditions (inaudibility) of technical aspects were challenges to listening tasks. Voice anxiety issues were common with oral performance tasks.

4.4.2. Teacher Interviews

Individual Solutions

The *Individual Solutions* theme demonstrates the individualized responses of teachers as they confronted the complication of the new assessment regulation. As detailed in Figure 9, the most common strategies were *Additional Support for Students* (n = 4), *Individualized Assessment* (n = 2), and *Flexible Scheduling* (n = 2).

Teacher Quotes:

- **T2:** “For students who were anxious or had speech difficulties, I tried to create a calm environment and even conducted their speaking exam separately, so they could feel more comfortable.”
- **T3:** “I give students extra practice and links to listening exercises, especially before exams, to help them get used to listening tasks.”
- **T5:** “I sometimes adjust speaking questions or give hints to students who struggle, focusing on their individual level rather than sticking rigidly to one format.”

T7: “I skip unnecessary parts of the textbook to create time for speaking and listening activities, spreading the tasks over multiple lessons.”

Additional Support for Students (n = 4) was the most common response, both emphasized as a reliable approach in utilizing additional examinations, providing flexible exam hours, giving either more practice, calming techniques, or further resources. *Individual Assessment* (n = 2) and *Flexible Scheduling* (n = 2) were less frequent responses but indicate efforts to work with the regulation in a way that would not overwhelm students while also achieving curriculum outcomes.

4.5. Chapter Summary

Using information from 89 learners and 7 language teachers, the chapter was structured around three research questions exploring participants' perceptions of the new assessment regulation, perceived advantages and disadvantages of the regulation, and the experience of implementing the regulation in classrooms. The data was inductively analyzed using MAXQDA software, and themes and codes were articulated to represent themes that were consistent across the responses. Both students and teachers recognized some communicative benefits of the regulation, particularly speaking and listening communicative benefits. At the same time there were genuine concerns which overpowered the perceived benefits regarding regulation's practicality, clarity and emotional impact. The themes noted students' individual learning benefits and anxieties, and teachers' reflections on the practical implementation and classroom influence which echoed with teachers' references to pedagogical possibilities and systemic constraints.

CHAPTER FIVE
DISCUSSION, CONCLUSION, AND IMPLICATIONS

5.1. Introduction

In this chapter, the interpretation and the discussion from the findings in regard to the research questions are presented. This chapter provides a detailed interpretation of how students and teachers experienced, reflected on, and evaluated the new regulation about assessment and evaluation of the English language. The analysis draws on thematic patterns, example responses, and relevant literature to demonstrate the results of this study.

5.2. Teacher and Student Perceptions and Reflections of the Regulation

The findings of this study suggest that both students and teachers tended to view the revised English language assessment regulation with a generally positive attitude. Yet, such perceptions were not uniform or uncritical; rather, they appeared to be mediated by the participants' respective institutional roles and the expectations placed upon them within the educational system. Students, for the most part, responded in terms of how the regulation influenced their personal learning and engagement with the language, whereas teachers adopted a broader evaluative lens, weighing pedagogical coherence against the practicalities of classroom implementation.

Among student participants, the theme *Improvement in Learning* emerged with the greatest frequency. This finding appears to indicate a shift in learners' orientation toward English, particularly in their increased willingness to invest time and attention in the subject. Several students reported that they had begun to study more regularly and with heightened focus. Although these responses were often brief and affectively framed, they point toward a form of positive washback which is a phenomenon wherein the structure of assessment begins to shape the nature of the learning process itself. Subcodes from teacher data such as *Speaking Skill* and *Listening Skill* further underscore a perceived enhancement in communicative competence. These observations resonate with long-standing arguments in the language assessment literature, particularly those advocating for alignment between

testing formats and authentic language use (see Canale and Swain, 1980; Bachman and Palmer, 1996). From this perspective, the students' accounts may be interpreted as evidence of constructive alignment between instructional aims and assessment design.

Teachers' responses, while similarly reflecting support for the new regulation, displayed a more nuanced and at times ambivalent tone. The dominant theme *Conceptually Positive* indicated an overall appreciation for the reform's communicative orientation. Some participants specifically acknowledged that the move away from grammar-dominated assessment toward tasks involving real-world language use marked a desirable pedagogical shift. However, this conceptual endorsement was frequently accompanied and dominated by expressions of concern, particularly regarding the policy's uneven implementation and the perceived lack of institutional infrastructure that is required to support its success. This tension recalls what Looney (2011) has described as a *policy-practice gap*, wherein the theoretical promise of a reform fails to materialize in everyday teaching contexts. Previous research has drawn attention to the complexity of educational change, highlighting the ways in which even well-intentioned reforms may falter in the absence of coherent systemic support (Cohen, 1990; Spillane et al., 2002). From this standpoint, the present data may serve evidence of the structural and contextual limitations that constrain meaningful pedagogical transformation.

As previously mentioned, educators' beliefs were clearly evident about this gap. Teachers believed there was a policy-practice gap despite perceiving the reform as C*onceptually Positive*. The reform was worthwhile conceptually, yet enacting this reform was often interrupted due to context by having large class sizes, time constraints, and limited other supporting structures. Rea-Dickins (2004) noted that teachers are both implementers of assessment policy, as well as interpreters of it. In relation to this, while educators can develop their own agency to bridge the policy-practice gap, their actions must also be supported by greater system level structures and resources. Even with positive intentions, policy reforms may appear superficial, shallow, or inconsistent. Therefore, the trend of perceived policy-practice gap in this study is not a manifestation of misunderstanding or resistance to assessment reform but exists as a structural challenge to resolving the potential of communicative assessment reform.

Interestingly, both groups agreed that students showed improvement in their speaking and listening skills, albeit from different perspectives. The students' perspective linked improvement to their perceptions of improved learning, skill development, and academic improvement. Teachers' observations were an additional point of reference and saw improvement based on student participation and output. This intersection highlights a shared perception of value, grounded in task authenticity and the principle that language is best learned and evaluated through use (Norris, 2016). Overall, teachers' observations were reflective and professional while students' comments were brief, emotional, and based on the immediacy of their classroom experience. It appears there is a divergence in the reflection of assessment based on the role of included stakeholders.

Positive sentiment prevailed overall; however, significant discrepancies were revealed. Importantly, while students refrained from outright rejecting the regulation concept, many expressed various mixed sentiments such as stress and/or confusion at the outset. Codes including *Stress and Anxiety*, *No Improvement*, and *Difficulty Understanding*, which reflect the psychological and cognitive workloads that students appeared to experience. Some students described feelings of anxiety related to the format of the speaking exam or expressed that they found listening difficult to understand. Supporting literature warns that test anxiety increases when performance-based assessment is implemented without adequate scaffolding (Hamp-Lyons, 2017; Cheng et al., 2008).

Conversely, teacher reactions tended to be more tempered. Even teachers who state that there were benefits to the reform provided qualifiers around the need for additional scaffolding structures. Teachers' concerns dealt not with the communicative nature of the policy, but instead with whether or not it could be executed across diverse populations. For example, some teachers mentioned that some students were simply not developmentally or linguistically ready for a speaking assessment. This view echoes what Green (2007) refers to as symbolic alignment reforms that appear progressive on paper, for example, but lack systemic capacity to be adapted in ways that have pedagogical meaning.

It is important to note that students rarely commented about their understandings in pedagogical terms, their accounts tended to be evaluative of the regulation in terms of personal outcomes ("my English improved", "I liked speaking more"), whereas teachers' comments were framed in terms of curriculum alignment, learner variability and constraints.

The difference seems to highlight the importance of self-evaluative aspects of divergent perspectives in research on educational reform, supporting Rea-Dickins (2004)'s statement on how the same policy is experienced in different roles or sectors.

Overall, the perceptions of both groups suggest that the new regulation is facilitating more and more meaningful learning experiences for students, if only because they are now actively engaging with the language in a wider range. However, the findings also uncovered areas of contention. For example, while students liked the change, they sometimes experienced affective or cognitive challenges; teachers were generally supportive of the reform yet felt that it needed to be backed up with structural support. As a result, student and teacher perspectives converge up to a point in principle, yet began to diverge with the realities of lived experience which complicates the process of enacting educational reform.

5.2.1. Positive Perceptions of the Regulation

The findings of this study indicate that while some students and teachers acknowledged certain benefits of the revised English language assessment regulation, their responses more frequently focused on its challenges. These perceptions were not uniformly positive and were often shaped by the participants' roles, expectations, and the constraints of their educational contexts. Students, for the most part, responded in terms of how the regulation influenced their personal learning and engagement with the language, whereas teachers adopted a broader evaluative lens, weighing pedagogical coherence against the practicalities of classroom implementation.

Among student participants (n = 89), the theme *Improvement in Learning* (f = 42) emerged with high frequency. This finding appears to indicate a shift in learners' orientation toward English, particularly in their increased willingness to invest time and attention in the subject. Several students reported that they had begun to study more regularly and with heightened focus. Although these responses were often brief and affectively framed, they point toward a form of positive washback, wherein the structure of assessment begins to shape the nature of the learning process itself. Subcodes such as *Speaking Skill* (f = 49) and *Listening Skill* (f = 20) further underscore a perceived enhancement in communicative

competence. These observations resonate with long-standing arguments in the language assessment literature, particularly those advocating for alignment between testing formats and authentic language use (see Canale and Swain, 1980; Bachman and Palmer, 1996). From this perspective, the students' accounts may be interpreted as evidence of constructive alignment between instructional aims and assessment design.

Teachers' responses, while occasionally reflecting support for specific aspects of the regulation, displayed a more nuanced and at times ambivalent tone. The theme *Conceptually Positive* indicated a degree of appreciation for the reform's communicative orientation. Some participants specifically acknowledged that the move away from grammar-dominated assessment toward tasks involving real-world language use marked a desirable pedagogical shift. However, this conceptual endorsement was frequently accompanied by expressions of concern, particularly regarding the policy's uneven implementation and the perceived lack of institutional infrastructure that is required to support its success. This tension recalls what Looney (2011) has described as a policy-practice gap, wherein the theoretical promise of a reform fails to materialize in everyday teaching contexts. Previous research has drawn attention to the complexity of educational change, highlighting the ways in which even well-intentioned reforms may falter in the absence of coherent systemic support (Cohen, 1990; Spillane et al., 2002). From this standpoint, the present data may serve as evidence of the structural and contextual limitations that constrain meaningful pedagogical transformation.

Both groups recognized some improvement in speaking and listening skills, but their interpretations of this improvement differed substantially. The students' perspective linked improvement to their perceptions of improved learning, skill development, and academic improvement. Teachers' observations were an additional point of reference and saw improvement based on student participation and output. This intersection highlights a shared perception of value, grounded in task authenticity and the principle that language is best learned and evaluated through use (Norris, 2016). Overall, teachers' observations were reflective and professional while students' comments were brief, emotional, and based on the immediacy of their classroom experience. This reflects the different lenses through which teachers and students interpret educational reforms.

Although positive experiences were present in the overall narrative especially among teachers, it was marked by significant concerns. While students did not reject the idea of the

regulation, many expressed mixed or conflicted responses involving stress, confusion, and practical difficulty. Codes including *Anxiety and Stress*, *No Improvement Felt*, and *Exam Format* reflect the psychological and cognitive workloads that students experienced. Some students described feelings of anxiety related to the format of the speaking exam or expressed that they found listening difficult to understand. Supporting literature warns that test anxiety increases when performance-based assessment is implemented without adequate scaffolding (Hamp-Lyons, 2017; Cheng et al., 2008).

Teachers' responses, however, leaned more critically overall, with only limited expressions of perceived success. Even teachers who stated that there were benefits to the reform provided qualifiers around the need for additional scaffolding structures. Teachers' concerns dealt not with the communicative nature of the policy, but instead with whether or not it could be executed across diverse populations. For example, some teachers mentioned that some students were simply not developmentally or linguistically ready for a speaking assessment. This view echoes what Green (2007) refers to as symbolic alignment reforms that appear progressive on paper but lack systemic capacity to be adapted in ways that have pedagogical meaning.

Conceptual Agreement and Implementation

The theme *Conceptually Positive* appeared in teacher responses as a sign of general agreement with the aims of the regulation. Several teachers stated that the shift toward communicative assessment aligned well with current pedagogical thinking, particularly in emphasizing skills that are often overlooked in grammar-based evaluation systems. T1, for instance, commented, "I believe the emphasis on speaking and listening is in line with real language use," while T4 mentioned that "the idea behind the new assessment is actually useful." These reflections indicate partial professional support for the overall idea behind the regulation.

However, this agreement remained at the conceptual level and was not always supported in practice. The positive views expressed were often accompanied by serious concerns regarding how the regulation had been implemented. Teachers frequently

mentioned issues such as the lack of clear rubrics, insufficient preparation time, or inadequate training, which prevented them from fully realizing the intended goals. As T6 explained, “I agree with the logic of focusing on communicative skills, but we didn’t get enough support to carry it out effectively.” This kind of cautious approval reflects what Looney (2011) describes as a disconnection between policy intent and institutional readiness.

Student responses did not include a direct equivalent to *Conceptually Positive*, but there were comments reflecting acceptance or understanding of the purpose of the regulation. A few students expressed that speaking and listening were useful skills to assess, though such views were embedded within broader statements about classroom experience or test anxiety. For example, S32 wrote, “I was nervous, but I know it helps with real English,” while S54 noted, “Listening is hard, but it makes sense because we need it in real life.” These comments suggest that although students did not explicitly endorse the regulation, some understood or accepted its intended purpose.

In general, teachers appeared more aware of the conceptual basis of the regulation, but their support depended on whether the system could be applied effectively. Students, in contrast, showed more fragmented or affective forms of agreement, often connected to classroom activities or self-perceived improvement. This distinction reinforces the need to evaluate both cognitive and emotional aspects of stakeholder alignment with educational reforms.

Perceived Benefits: Skill Development and Academic Improvements

The implementation of the new English assessment regulation had perceived benefits, particularly in relation to communicative skills and academic improvement. Student responses revealed *Speaking Skill Improvement*, *Improvement in Learning*, and *Listening Skill Improvement* as the most commonly mentioned advantages. These responses suggest that students were not only more exposed to speaking and listening tasks but also viewed them as meaningful steps in their learning.

Such perceived improvements are in agreement with (CLT) principles, which emphasize interaction and task-based instruction as essential for developing language

proficiency (Richards, 2006). Luoma (2004) also notes that continuous exposure to speaking tasks can increase fluency and lower students' hesitation towards speaking. In this context, when speaking and listening are embedded in routine assessment, learners may become more confident and engaged in using English functionally.

Beyond speaking skills, some students described academic improvements regarding motivation and independent learning. The codes *Improved Grades*, *Memorization and Retention*, and *Enhanced Self-Study Habits* reflected this. S10 stated, "The new exams made me study by myself and focus more on improving my vocabulary." This kind of reflection connects to learner autonomy, on which Little (2007) comments that students engage more when assessments are seen as purposeful and authentic.

Teachers also acknowledged some of these gains, though in more restrained ways. The theme *Improved Student Skills* showed that some educators observed stronger classroom participation and increased speaking attempts among previously withdrawn students. T3 noted, "Even a student who has never spoken before can discover themselves during the speaking exam." Such feedback points not only to skill development but also to shifts in learner self-perception, which are considered valuable in communicative assessment (Bygate, 2009).

However, the data for *Improved Grades* also revealed complexity. Although students expressed satisfaction with higher scores, many of these comments co-occurred with stress or learning difficulties. S23 reported, "My grades in English have noticeably improved," but also acknowledged challenges with listening tasks. Similarly, S24 mentioned, "It helped me get higher grades," yet added, "I get stressed when I am studying for this speaking exam." These examples show that rising scores did not always reflect deeper learning and may, at times, represent surface-level achievement driven by exam strategies rather than communicative growth.

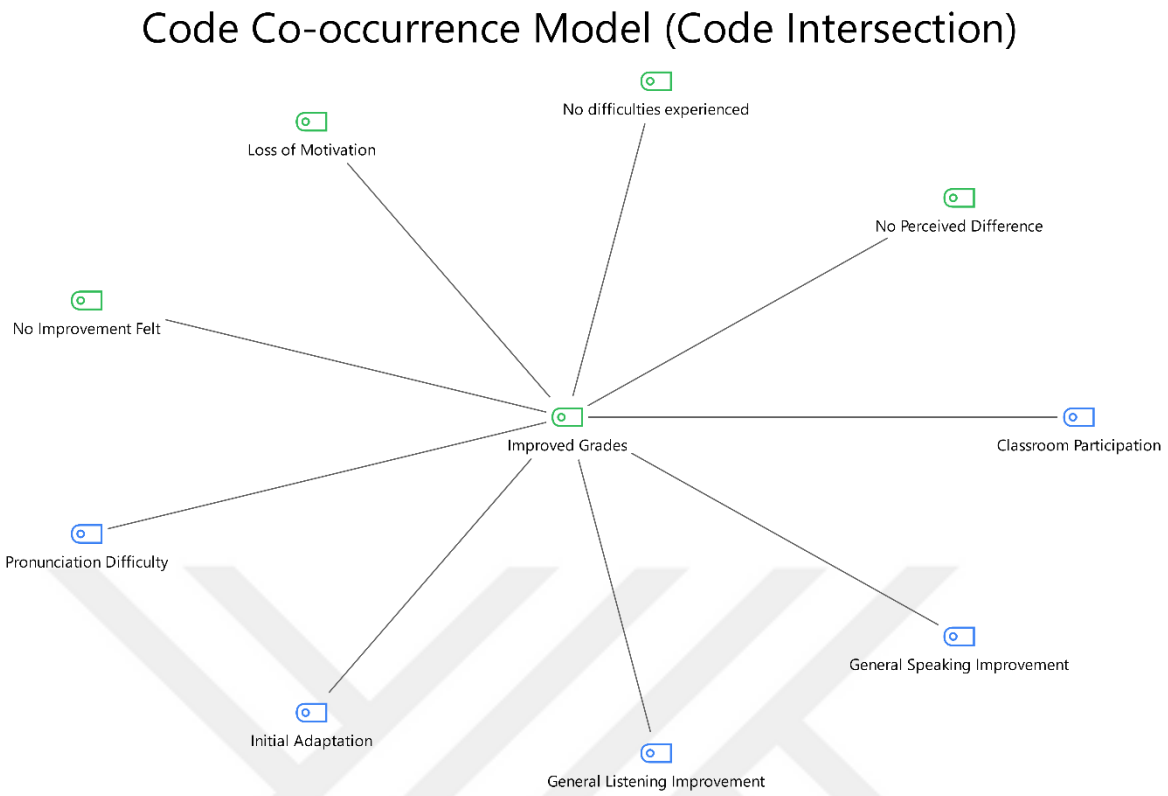

Figure 14. MAXQDA code co-occurrence map of Improved Grades

This pattern is also visible in the code co-occurrence model (see Figure 2), where *Improved Grades* frequently intersects with emotionally and cognitively challenging codes such as *Speaking Anxiety*, *Pronunciation Difficulty*, and *No Improvement Felt*. On another note, the code *Improved Grades* co-occurred more often with negative or stress-related codes than with other improvement-focused ones, suggesting that students who reported improved grades may have simultaneously experienced emotional strain or questioned their actual learning. This raises concerns that while grades improved, students may have prioritized exam success over genuine language development. It also suggests that the regulation may be shaping a test-driven environment where grades are valued more than communicative competence.

This reflects the classic washback effect, where test design shapes not only classroom practices but also students' emotional engagement (Cheng, 2005; Tsagari, 2009). If students' motivation is primarily tied to grades, rather than growth, then the policy may be reinforcing

exam-centered learning at the cost of communicative competence. As one student bluntly noted, “It raised my grades, but when I mispronounce words, I get confused” (S42).

Although some students reported increased self-study and vocabulary focus, most comments related to grades were also linked with stress, limited skill development, or a sense of no meaningful change. These findings suggest that progress and improvements under the new system may be felt in categories of grades and skills, but not always in confidence or competence. Without addressing the emotional and pedagogical support needed, higher performance may remain a surface achievement rather than a sign of meaningful learning.

5.2.2. Negative and Mixed Perceptions and Affective Responses

Negative or mixed perceptions appeared frequently across both student and teacher data. Among student participants, themes such as *Anxiety and Stress*, *Assessment-Related Challenges*, and *No Improvement Felt* ($n = 17$) occurred with high frequency. Many students expressed feelings of pressure and emotional discomfort related to the new format. S64 stated, “With the new system I started to struggle and feel stressed,” and S14 mentioned, “I get nervous when I have to speak in front of the class; it makes me forget what I want to say.”

Beyond emotional responses, students also noted practical issues that made the process difficult. Listening tasks were often reported as too fast or technically unclear. S11 commented, “Conversations in listening tracks are too fast,” while S8 explained, “Listening tracks were not audible on smartboard.” These comments show that negative perceptions were not limited to affective reactions but extended to task design and delivery.

Some students also questioned the learning impact of the new assessment system. Although improvements in grades were mentioned by several, others were unconvinced about any real learning gain. For example, S60 noted, “I don’t think the new exams made any difference in my learning,” and S61 commented, “Its influence was not good; lessons are way more difficult now.” This indicates that academic benefit was not universally experienced or acknowledged.

Teachers also expressed a predominantly critical stance, with multiple references to implementation problems. Themes such as *Time Constraints*, *Unclear Assessment Criteria*, and *Student Anxiety and Stress* reflect broad dissatisfaction. T2 reported, "The lack of rubric and time limited our ability to prepare students," while T7 mentioned, "Speaking exams caused panic in students and affected classroom atmosphere."

Several teachers raised equity concerns as well. The system was seen as more challenging for students with lower proficiency levels or emotional difficulties. T6 noted, "Some students are not mentally ready for this kind of assessment, and that creates unfairness," suggesting that the policy may have unintentionally increased disparities in the classroom. This aligns with research by Green (2007), who warned against reform measures that appear inclusive in design but result in selective success in practice.

In summary, both groups voiced significant challenges related to emotional pressure, unclear expectations, and inequitable experiences. Although some acknowledged isolated improvements, the overall tone of these responses leaned toward concern rather than support. These findings point to the importance of clearer guidance, emotional scaffolding, and context-sensitive implementation when designing assessment reform.

5.3. Reported Limitations and Implemented Solutions

While the regulation is based on communicative and skills-based goals, many students reported emotional challenges and limitations during its application. One of the most frequently mentioned issues was *Anxiety and Stress* ($n = 27$; $f = 28$), closely followed by *Assessment-Related Challenges* ($n = 24$; $f = 27$) and *Difficulty with Listening Tasks* ($n = 22$). These patterns highlight a contradiction in the policy's intended versus experienced outcomes.

Student reflections showed emotional discomfort. S14 admitted, "I get nervous when I have to speak in front of the class; it makes me forget what I want to say," and S64 added, "With the new system I started to struggle and feel stressed." These accounts suggest that stress caused by speaking assessment may limit students' ability to learn effectively.

Krashen's Affective Filter Hypothesis (1982) supports this idea, explaining that emotional barriers such as anxiety and fear can block language input. When students perceive tasks as judgmental or intimidating, especially oral ones, they may shut down or freeze. As Horwitz (2001) notes, such environments make meaningful language engagement difficult.

Students also mentioned struggling with *Difficulty in Adaptation* (n = 10) and *Inhibited Learning* (n = 2; f = 2). Switching to a performance-based system disrupted familiar classroom routines. According to MacIntyre and Gardner (1991), repeated anxiety may lead to long-term reluctance in language learners and if the support is lacking, learners may eventually avoid speaking tasks or perform well below their potential.

Stress was also tied to academic outcomes. S42 stated, "My grades went down because I'm not good at listening exams," while S56 commented, "The new system made me feel less motivated since I am always stressed about speaking." These reactions reflect the washback effect described by Tsagari (2009), where test format influences both learning and emotions.

The goal of the regulation was to foster meaningful communication. However, stress and pressure may have reduced its effectiveness. Some solutions could include emotional scaffolding training for teachers, rethinking assessment formats, or adding strategies to ease classroom tension.

Some teachers did take action to soften these effects. Codes under *Implemented Solutions* showed efforts like *Providing Support* (n = 4), *Individual Assessment* (n = 2), and *Flexible Scheduling* (n = 2). T5 explained, "I give students extra practice to reduce anxiety before the oral part." T6 added, "I let students take the oral part individually if they are too anxious to perform in front of the class." These responses suggest that despite time or system limitations, some teachers worked to create a more supportive classroom.

5.3.1. Institutional and Pedagogical Barriers in Practice

Teacher interviews revealed a range of school- and teaching-related barriers that made applying the new assessment system difficult. The results showed that *Time Constraints* (n = 6) was the most frequent code in terms of such obstacles mentioned for teachers, followed by *Unclear Assessment Criteria* (n = 5) and *Curriculum Misalignment* (n = 4). These barriers further illustrate the ongoing misalignment between the objectives of the policy and real-world classroom experience.

Teachers consistently pointed out that their class hours were already too limited to implement communicative language teaching, and the additional planning, conducting, and assessing of speaking and listening exams placed even more pressure on their time. T1 stated, "We already have very limited hours, and now we need extra time for speaking exams, which disrupts the schedule." T2 echoed this concern, adding, "Sometimes we have to skip parts of the regular curriculum just to make room for these assessments."

Another teacher (T3) mentioned the lack of a standardized rubric: "There is no clear rubric for assessing speaking; we all grade differently," which revealed ambiguity and inconsistency in teachers' assessment. T4 further noted, "Everyone is doing their own thing. I don't know what other teachers use, so there's no shared model."

Moreover, T5 stated, "The curriculum is not made for communicative assessments, and we have to improvise many times," indicating the misalignment between the pedagogical content of their instruction and assessment. This led to lessons that felt disconnected from the types of performance expected in the new exams.

Together, the previous statements pointed to literature that highlights the gap between policy and practice, whereby a common occurrence in reforming education is the presence of bold policy without support structures, professional development, or distributive resources (Datnow, 2020; Spillane, 2004). Additionally, policies that aim to build teacher agency and overlook the practice of the classroom context may result in more demands on teachers, given that the implementation is not systemic or collaborative. Fullan (2007) argued that failure to incorporate teacher agency and practical classroom realities into policy causes superficial or inconsistent policy implementation in practice.

In addition, the research suggests that a successful policy will need to address alignment not just in the structures of curriculum, pedagogy, and assessment, but the capacity of teachers and system (Tao and Gao, 2017). Within the Turkish context, these challenges were very much present as teachers had to “improvise” to fit communicatively assessed tasks into system modes that were still strongly focused on traditional aspects of written and grammar-based assessments. This gap shows that although the intention was to improve communicative competence, structural limitations prevented it from being applied in everyday teaching.

The magnitude of these challenges implies that policy change alone is not enough, and as Shohamy (2017) notes, political and institutional power structures must be addressed when examining language assessment change. In the absence of comprehensive provision and time to provide change instruction through systematic and thoughtful restructuring, including training resources and manageable workloads, teachers may conceptually agree with the reform, but not practically.

5.3.2. Adaptive Strategies and Solutions

As teachers faced an array of challenges brought on by the new regulation, teacher data surfaced a noteworthy tendency toward Individual Solutions to address barriers to implementation. These teacher responses were coded under the category Individual Solutions and included subcodes such as Student Improvement (n = 4; f = 12), Individual Assessment (n = 2; f = 3), and Flexible Scheduling (n = 2; f = 3). These frequencies indicate that while not necessarily implemented universally, several teachers did seek ways to implement the policy in a manner that lowered anxiety for themselves, while also increasing access to students.

Teachers offered a number of strategies tailored to their classroom realities that supported student participation in the performance-based assessments and allowed students to feel more confident and capable. For instance, one teacher explained, "I tried to make the environment calm, and I even did their speaking exam separately, so they would feel more comfortable" (T2). Another teacher stated she provided her students with extra practice and

listening exercises to help them prepare for the listening task (T3). Some provided different levels of speaking questions based on students' proficiency levels, and some also skipped some of the more relevant textbook content to allow them more time for more communicative type tasks. T5 mentions adjustments as "I sometimes adjust speaking questions or give hints to students who struggle, focusing on their individual level rather than sticking rigidly to one format", and T7 comments ""I skip unnecessary parts of the textbook to create time for speaking and listening activities, spreading the tasks over multiple lessons" on having to skip content.

These examples are indicative of an awareness of instructional differentiation, diminished anxiety, and affective support for students in relation to assessment. While still limited in scope, such evolving assessment strategies also speak to the influence of teacher agency when responding to top-down policy mandates. These practices align with what Carless (2012) refers to as "assessment adaptation," where practitioners adjust their implementation to the needs of students and, in doing so, bridge potential gaps between policy expectations and classroom realities.

There may be a number of mediating factors that contributed to this lack of student-sourced solutions. First, it may be that students do not have the assessment literacy or experience it takes to imagine alternative forms of implementation. Second, students' stress and anxiety about performance, as well as perceived unfairness as stated in their comments, "It would be better if there were not speaking and listening tests" (S5), or "Grading system was difficult for me because sometimes it felt unfair" (S15), took precedence over their ability to begin to think critically about productive improvement. In those cases, learners tend to focus more on discomfort and less on pedagogical changes, particularly when they are not included in the assessment process themselves (Schaefer et al., 2024).

Third, this may also reflect findings in educational research that identify students frequently feel disempowered or excluded from decision-making when it comes to assessment policy and practice (Cook-Sather, 2006). Within a centralized education system such as Türkiye's, students may not perceive themselves as active participants who can influence assessment design. As a result, they may adopt a disempowered stance even when invited to provide feedback.

As a collective finding, the data reveals a striking asymmetry: some teachers actively work to overcome the limitations of the system using adaptive practices, whereas student voice primarily emerges in the form of frustration, yet without accompanying proposals for change. This indicates that future implementations should also consider student voice and feedback mechanisms, so learners can participate meaningfully in shaping assessment practice. This is in line with studies that advocate for participatory engagement in which students' voices play a role in guiding educational decisions (Schaefer et al., 2024).

In summary, the Individual Solutions that teachers identified were local strategies to balance institutional expectations with students' needs. If policies are to be effectively constructed in the future, it is critical to establish mechanisms that cultivate agency and ownership among both teachers and learners. This would not only enhance reform implementation but also help ensure sustainability and relevance in the broader educational landscape.

5.4. Conclusion

This research investigated how the recent English assessment regulation implemented by the Turkish MoNE in 2023 was perceived and experienced by English teachers and students in a middle school context. The new regulation places equal value on communicative skills, including listening and speaking, in conjunction with more traditional written assessments. This represents a significant policy change in the Turkish EFL context. This qualitative research gathered, coded, and analyzed interview data from 7 English teachers and open-ended survey data from 89 seventh graders, in order to examine the perceptions of these two groups of stakeholders, the perceived advantages and disadvantages of the new assessment system, and the challenges each group experienced during the implementation process. This study sheds light on the complexities of how policy is enacted in practice, revealing the ways in which intended reforms are influenced by storied realities of emotional, institutional, and pedagogical nature in classrooms.

Students and teachers offered important insights into how the new regulation is perceived. Many students found the speaking and listening components helpful for real-life

communication, reporting greater engagement and practice opportunities. Teachers largely supported the idea behind the reform, particularly the focus on neglected skills, but voiced hesitation about full implementation. Teachers mostly talked about the lack of time, unclear grading rules, and having too many students to manage well. This shows a clear gap between agreeing with the idea of the reform and being able to apply it without problems.

Students said they improved in speaking and listening. Some also felt they did better in school because of this. Their positive reports support the reform's aim of boosting communicative competence (Canale and Swain, 1980). Teachers saw that students had joined in more, but they also noticed emotional struggles. Many students said they were nervous during oral exams. Teachers also recognized this pressure. This connects with Krashen's (1982) idea that stress can get in the way of using a new language. It's important to prepare students not just with language lessons but also with ways to manage stress in speaking and listening tasks.

Teachers struggled with daily issues like not having enough time, unclear rubrics, and a mismatch between lessons and exams. Despite these challenges, a few teachers found their own solutions, like giving exams at different times or assessing students one by one. These practices demonstrate teacher agency and suggest that practical workarounds can ease the impact of rigid systems. Student responses, in contrast, focused mainly on their emotional experiences and performance concerns. Their feedback highlights the need for reforms to consider affective factors and to involve learners more meaningfully in shaping classroom assessment.

In summary, the findings suggest that the success of an assessment reform cannot be measured solely in terms of its embrace of progressive pedagogical ideals. The MoNE regulation exemplifies a move toward the integration of skills, which is espoused within the tenets of communicative language teaching. However, it is the extent to which these can be merged with other realized realities in the classroom, context and stakeholder, that ultimately convey relevance. Emotional readiness, logistical possibility and professional capacity are just as integral to reform outcomes as curricular content. This notion aligns with scholars such as Shohamy (2017) asserts that assessment policy and systems are not neutral; they have ideological and operational implications that can directly affect both educators and learners.

Involving teachers and students in the process of implementation and design of the educational reforms is crucial to the effectiveness of implementing it. For a reform to be successful, schools must prepare teachers, provide adequate allocation in terms of resources, and offer flexibility in implementation and allow an honest, contributory system that stakeholders are able to share ideas. Without such efforts, even the most pedagogically aligned reforms run the risk of failing or becoming counterproductive.

This study prioritizes stakeholder viewpoints and experiences. It also lays the groundwork for future research on the long-term impact of such reforms, comparative studies across different regions or student levels, and the development of assessment tools that prioritize both validity and learner well-being.

5.5. Implications and Suggestions

This study shows that for assessment reforms to succeed, teacher conditions and student needs must be addressed as a whole. The new regulation brought useful goals, like more attention to speaking and listening, but also came with problems that made it hard to apply in practice. Teachers and students both commented that policy only works when it comes with adequate support, in addition to the rules.

One clear implication concerns the curriculum. There is a mismatch between the fast pace of the curriculum and the time needed for skill-based teaching. Reducing unit load and making space for communicative tasks would make it easier for teachers to apply the policy. More frequent English lessons during the week could give teachers time to plan, practice, and assess speaking and listening without rushing.

Another key concern was the lack of shared and centralized grading tools. Teachers emphasized the need for clearer and more unified assessment criteria, especially for speaking exams. Even if full standardization is not possible, a shared structure could help reduce confusion and improve fairness. Along with this, teachers asked for yearly plans that include dedicated space for speaking and listening assessments, so that these tasks are no longer added at the last minute or done under pressure.

Teachers also noted that current textbooks are not designed to support the new speaking and listening goals. Any future reform should be planned together with updated instructional materials that support the skills being tested. Otherwise, teachers are left to "fill the gap" between what the curriculum offers and what the exams expect.

Many teachers said the regulation came too fast. This caused stress, confusion, and extra workload. They were not trained, and some felt confused or overwhelmed. Before introducing any future changes, it is essential to conduct field research, listen to teachers' voices, and provide training in advance. Without preparation, even the best-looking policy may not work in real classrooms.

On the student side, some important emotional risks appeared. Students frequently mentioned feeling stressed before the speaking exam. Others reported that they were more focused on grades than on learning, and some felt overwhelmed by the overall increase in exam load. This kind of stress can lower motivation and hurt the learning process itself. At the same time, students also said they found the new system helpful for improving their speaking and listening, which suggests that the core idea behind the reform has value. The challenge is to make that idea work in a way that feels safe and fair.

Based on these findings, this study offers the following suggestions:

- Reduce the number of units in the curriculum to make space for communicative tasks.
- Increase the number of English class hours in the weekly schedule.
- Create a shared but flexible speaking assessment framework for all schools.
- Update textbooks so they match the new regulation and include more speaking-focused content.
- Include speaking and listening exams in annual lesson plans from the beginning of the year.
- Provide clear communication, in-service training, and enough time before any major reform is introduced.

- Design low-stakes speaking opportunities that help students build confidence before formal exams.

With better support, both students and teachers may feel more ready, more included, and more able to succeed.

5.6. Chapter Summary

In this chapter, the study's findings were brought together and connected to what they signify for classroom practice. Useful information was provided by both students and teachers in light of the English assessment regulation brought by MoNE in 2023. Even though many expressed support for a focus on speaking and listening, there were also a number of barriers. Some of those barriers included pressure, grading ambiguity, confusion about results, lack of time, and materials that did not adhere to the new expectations.

The dialogue also highlighted that the majority of teachers endorsed the changes in principle but faced difficulty understanding or interpreting the changes in practice. To help teachers adjust to the changes in practice, they suggested appropriate policy ideas such as a reduction in the number of units, provision of more class time, and creation of shared assessments to support this change in practice. Students reported that the exam process resulted in them becoming better speakers, even though many felt anxious and over-focused on grading. The feedback indicated that emotional support is equally significant as a meaningful measure of progress.

The chapter also outlined possible methods of moving forward, including opportunities for teacher development, new resources (e.g. books), improved planning time, and detailed research before implementing future regulations and reforms with appropriate time and support, both teachers and learners are able to apply educational policy into classroom practices.

REFERENCES

Acar-Erdol, T., and Yıldızlı, H. (2018). Classroom assessment practices of teachers in Turkey. *International Journal of Instruction, 11*(3), 587-602. https://doi.org/10.12973/iji.2018.11340a

Alderson, J. C. (2000). *Assessing reading*. Cambridge University Press.

Alderson, J. C., Clapham, C., and Wall, D. (1995). *Language test construction and evaluation*. Cambridge University Press.

Alderson, J. C., and Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing, 13*(3), 280–297. https://doi.org/10.1177/026553229601300305

Al-Garni, S. A., and Almuhammadi, A. H. (2019). The effect of using communicative language teaching activities on EFL students' speaking skills at the University of Jeddah. *English Language Teaching*, *12*(6), 74–80. https://doi.org/10.5539/elt.v12n6p74

Ansari, M. S. (2015). Speaking anxiety in ESL/EFL classrooms: A holistic approach and practical study. *International Journal of Educational Investigations*, *2*, 38-46.

Bachman, L. F. (1990). *Fundemental considerations in language testing.* Oxford University Press.

Bachman, H. D. (2004). *Statistical analyses for language assessment*. Cambridge University Press.

Bachman, L. F., and Palmer, A.S. (1996) *Language testing in practices*. Pearson Education.

Bayyurt, Y. (2008). A lingua franca or an international language: The status of English in Turkey. In *ELF Forum, Helsinki, Finland*.

Black, P. and William, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, *5*(1), 7-74. http://dx.doi.org/10.1080/0969595980050102

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brindley, G. (1998). Assessing listening abilities. *Annual Review of Applied Linguistics*, *18*, 171–191. https://doi.org/10.1017/s0267190500003536

Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Pearson Education.

Brown, H. D., and Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices* (2nd ed.). Pearson Education.

Brunfaut, T. (2016). Assessing listening. In. D. Tsagari and J. Banerjee (Eds.), *Handbook of Second Language Assessment* (pp. 97-112). De Gruyter Mouton. https://doi.org/10.1515/9781614513827-009

Buck, G. (2001). *Assessing listening*. Cambridge University Press.

Burns, A., and Richards, J. C. (Eds.). (2012). *The Cambridge guide to pedagogy and practice in second language teaching*. Cambridge University Press.

Bygate, M. (2009). Teaching and testing speaking. In M. H. Long and C. J. Doughty (Eds.), *The handbook of language teaching* (pp. 412–440). Wiley-Blackwell. https://doi.org/10.1002/9781444315783.ch24

Byram, M. (2021). *Teaching and assessing intercultural communicative competence: Revisited* (2nd ed.). Multilingual Matters.

Calderón Procel, P. (2024). *Communicative approach and learner-centered classroom practices in EFL contexts* [Bachelor's thesis, Universidad Nacional de Educación (UNAE)]. UNAE Institutional Repository.

Canagarajah, S. (2006). Changing communicative needs, revised assessment objectives: Testing English as an international language. *Language Assessment Quarterly*, *3*(3), 229–242. https://doi.org/10.1207/s15434311laq0303_1

Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. Richards and R. Schmidt (Eds.), *Language and Communication*. Longman.

Canale, M., and Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, *1*(1), 1–47. https://doi.org/10.1093/applin/i.1.1

Carless, D. (2012). *From testing to productive student learning: Implementing formative assessment in Confucian-heritage settings*. Routledge

Cazden, C. B. (2011). Dell Hymes' construct of "Communicative Competence". *Anthropology and Education Quarterly*, *42*(4), 364-369, https://www.jstor.org/stable/41410139

Celce-Murcia, M. (2008). Rethinking the role of communicative competence in language teaching. In E. A. Soler and M. S. Jordà (Eds.) *Intercultural Language Use and Language Learning*. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-5639-0_3

Chan, S. H. (2018). Affective responses in oral proficiency assessment: A review. *Asian EFL Journal, 20*(12), 68–90.

Chapelle, C. A., and Voss, E. (2011). *Computer applications in applied linguistics*. Wiley-Blackwell.

Cheng, L. (2005). *Changing language teaching through language testing: A washback study*. Cambridge University Press.

Cheng, L, Rogers, W., and Hu, H. (2007). ESL/EFL Instructors' beliefs about assessment and evaluation. *Comparative and International Education*, *36*(1). https://doi.org/10.5206/cie-eci.v36i1.9088

Cheng, L., Watanabe, Y., and Curtis, A. (2008). *Washback in language testing: Research contexts and methods*. Routledge.

Cheung, K. K. C., and Tai, K. W. H. (2021). The use of intercoder reliability in qualitative interview data analysis in science education. *Research in Science and Technological Education, 41*(3), 1155–1175. https://doi.org/10.1080/02635143.2021.1993179

Cloudia Ho, Y.-Y. (2020). Communicative language teaching and English as a foreign language undergraduates' communicative competence in tourism English. *Journal of Hospitality, Leisure, Sport and Tourism Education*, *27*. https://doi.org/10.1016/j.jhlste.2020.100271

Cogo, A., and Dewey, M. (2012). Efficiency in ELF communication: From pragmatic motives to lexicogrammatical innovation. *Nordic Journal of English Studies, 5*(2), 59–94

Cook-Sather, A. (2006). Sound, presence, and power: "Student voice" in educational research and reform. *Curriculum Inquiry, 36*(4), 359–390. https://doi.org/10.1111/j.1467-873X.2006.00363.x

Coombe, C., Folse, K. S., and Hubley, N. (2007). *A Practical Guide to Assessing English Language Learners*. University of Michigan Press.

Coombe, C., Davidson, P., and O'Sullivan, B. (2020). *The Cambridge guide to second language assessment*. Cambridge University Press. https://doi.org/10.1017/9781139024390

Creswell, J. W., and Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.

Cummins, J., and Davison, C. (2007). International Handbook of English Language Teaching. Springer.

Das, A., Kuyini, A. B., and Desai, I. (2014). Inclusive education in India: Are the teachers prepared? *International Journal of Special Education, 29*(2), 1–16.

Datnow, A. (2020). The role of teachers in educational reform: A review of the research. *Educational Administration Quarterly, 56*(3), 452–478.

Denzin, N. K., and Lincoln, Y. S. (2011). *The SAGE handbook of qualitative research* (4th ed.). SAGE Publications.

Dinçay, T. (2010). A quick chronological review of the ELT methods along with their techniques and principles choosing eclecticism from among language teaching methods. *Dil Dergisi*, 40–62. https://doi.org/10.1501/dilder_0000000125

Erlingsson, C., and Brysiewicz, P. (2017). A hands-on guide to doing content analysis. *African Journal of Emergency Medicine, 7*(3), 93–99. https://doi.org/10.1016/j.afjem.2017.08.001

Field, J. (2009). *Listening in the language classroom*. Cambridge University Press.

Fulcher, G. (2000). The 'communicative' legacy in language testing. *System*, *28*(4), 483–497. https://doi.org/10.1016/s0346-251x(00)00033-6

Fulcher, G. (2003). *Testing second language speaking*. Pearson Education.

Fulcher, G., and Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. Routledge.

Fullan, M. (2007). *The new meaning of educational change*. Teachers College Press.

Genesee, F, and Upshur, J. A. (1996). *Classroom-based evaluation in second language education*. Cambridge University Press.

Gipps, C. V. (2012). *Beyond testing: Towards a theory of educational assessment*. Routledge.

Göktürk Sağlam, A. L. (2018). Can exams change how and what teachers teach? Investigating the washback effect of a university English language proficiency test in the Turkish context. *Eurasian Journal of Applied Linguistics*, *4*(2), 155–176. https://doi.org/10.32601/ejal.464094

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge University Press.

Green, A. (2006). Washback to the learner: Learner and teacher perspectives on IELTS preparation course expectations and outcomes. *Assessing Writing*, *11*(2), 113–134. https://doi.org/10.1016/j.asw.2006.07.002

Gronlund, N. E. (2006). *Assessment of student achievement* (9th ed.). Pearson Education.

Hamp-Lyons, L. (2017). Language assessment literacy for language learning-oriented assessment. *Studies in Language Assessment*, 88–111. https://doi.org/10.58379/lixl1198

Harding, L., and Kremmel, B. (2016). Assessment literacy and language teachers: Insights from the UK. *Language Testing, 33*(3), 327–345

Harlen, W., and James, M. (1997) Assessment and learning: Differences and relationships between formative and summative assessment. *Asesssment in Education: Principles, Policy and Practice*, *4*(3), 365-379. https://doi.org/10.1080/0969594970040304

Hattie, J., and Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hondo, J. (2018). Functional syllabus. *The TESOL Encyclopedia of English Language Teaching*, 1–13. https://doi.org/10.1002/9781118784235.eelt0689

Horwitz, E. K. (2001). Language anxiety and achievement. *Annual Review of Applied Linguistics, 21*, 112–126. https://doi.org/10.1017/S0267190501000071

Hsieh, H.-F., and Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*(9), 1277–1288. https://doi.org/10.1177/1049732305276687

Hughes, A. (2003). *Testing for language teachers*. Cambridge University Press.

Hunter, D., and Smith, R. (2012). Unpackaging the past: 'CLT' through ELTJ keywords. *ELT Journal*, *66*(4), 430–439. https://doi.org/10.1093/elt/ccs036

Hyland, K. (2003). *Second language writing*. Cambridge University Press.

Hymes, D. H. (1972). On communicative competence. In J. B. Pride and J. Holmes (Eds.), *Sociolinguistics: Selected readings* (pp. 269–293). Penguin.

Inamov, U., and Holmatova, M. (2025). Communicative language teaching (CLT) approach in language teaching. *Web of Teachers: Inderscience Research*, *3*(1), 351-353. https://webofjournals.com/index.php/1/article/view/3042

İnceçay, G., and Akyel, A. S. (2014). Turkish EFL teachers' perceptions of English as a lingua franca. *Turkish Online Journal of Qualitative Inquiry*, *5*(1). https://doi.org/10.17569/tojqi.84118

Jacobs, G. M., and Farrell, T. S. (2003). Understanding and implementing the CLT (communicative language teaching) paradigm. *RELC Journal*, *34*(1), 5–30. https://doi.org/10.1177/003368820303400102

Jenkins, J. (2009). English as a lingua franca: Interpretations and attitudes. *World Englishes*, *28*(2), 200–207. https://doi.org/10.1111/j.1467-971x.2009.01582.x

Kaçar, K. (2023). Commonly used techniques in testing foreign language skills and communicative language teaching. *International Journal of Field Education Researches, 4*(2), 68–85. https://doi.org/10.54971/synergy.1381532

Kırkgöz, Y. (2008). Globalization and english language policy in Turkey: Policy changes and implementations. *Educational Policy*, *23*(5), 663–684. https://doi.org/10.1177/0895904808316319

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.

Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174.

Liao, L., Ye, S. X., and Yang, J. (2023). A mini review of Communicative Language Testing. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1058411

Lichtman, M. (2023). *Qualitative Research in Education: A User's Guide* (3rd ed.). SAGE Publications.

Little, D. (2007). Language learner autonomy: Some fundamental considerations revisited. *Innovation in Language Learning and Teaching, 1*(1), 14–29. https://doi.org/10.2167/illt040.0

Littlewood, W. (2013). Developing a context-sensitive pedagogy for communication-oriented language teaching. *English Teaching*, *68*(3), 3–25. https://doi.org/10.15858/engtea.68.3.201309.3

Lincoln, Y. S., and Guba, E. G. (1985). *Naturalistic inquiry.* Sage Publications.

Looney, J. (2011). Developing high-quality teachers: Teacher evaluation for improvement. *European Journal of Education*, *46*(4), 440–455. https://doi.org/10.1111/j.1465-3435.2011.01492.x

Luoma, S. (2004). *Assessing speaking*. Cambridge University Press.

Ma, T. (2009). On communicative language teaching — theoretical foundations and principles. *Asian Social Science*, *5*(4). https://doi.org/10.5539/ass.v5n4p40

Mangaleswaran, S., and Aziz, A. A. (2019). The impact of the implementation of CLT on students' speaking skills. *International Journal of Scientific and Research Publications (IJSRP)*, *9*(4). https://doi.org/10.29322/ijsrp.9.04.2019.p8814

Mansory, M. (2020). The significance of non-traditional and alternative assessment in English language teaching: Evidence from literature. *International Journal of Linguistics*, *12*(5), 210. https://doi.org/10.5296/ijl.v12i5.17782

MacIntyre, P. D., and Gardner, R. C. (1991). Methods and results in the study of anxiety and language learning: A review of the literature. *Language Learning, 41*(1), 85–117. https://doi.org/10.1111/j.1467-1770.1991.tb00677.x

McNamara, T. (2017). *Language testing*. Oxford University Press.

Messick, S. (1996). Validity and Washback in language testing. *ETS Research Report Series*, *1996*(1). https://doi.org/10.1002/j.2333-8504.1996.tb01695.x

Ministry of National Education (2023). *Milli Eğitim Bakanlığı Ölçme ve Değerlendirme Yönetmeliği* [Regulation on assessment and evaluation by Ministry of National Education]. Resmi Gazete. https://www.resmigazete.gov.tr/eskiler/2023/09/20230909-2.htm

Nergis, A. (2011). Foreign language teacher education in Turkey: A historical overview. *Procedia - Social and Behavioral Sciences*, *15*, 181–185. https://doi.org/10.1016/j.sbspro.2011.03.070

Norris, J. M. (2016). Current uses for task-based language assessment. *Annual Review of Applied Linguistics, 36*, 230–244. https://doi.org/10.1017/S0267190516000027

Nowell, L. S., Norris, J. M., White, D. E., and Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods, 16*(1), 1–13. https://doi.org/10.1177/1609406917733847

Oktay, A. (2015). Foreign language teaching: A problem in Turkish education. *Procedia - Social and Behavioral Sciences*, *174*, 584–593. https://doi.org/10.1016/j.sbspro.2015.01.587

O'Sullivan, B. (2000). Exploring gender issues in language testing. *System*, *28*(4), 523-536.

Öztürk, G., and Gürbüz, N. (2014). Speaking anxiety among Turkish EFL learners: The case at a state university. *Journal of Language and Linguistic Studies*, *10*(1), 1-17.

Paker, T. (2012). Yabancı dil eğitimi ve öğretiminde programların tüm süreçlerinin denetimi ve bireysel yönlendirmenin rolü. In Hacettepe Üniversitesi Eğitim Fakültesi İngiliz Dili Eğitimi Anabilim Dalı (Ed.), *1. Yabancı Dil Eğitimi Çalıştayı Bildirileri (12–13 Kasım 2012)* (pp. 139–149). Hacettepe University Press. https://www.researchgate.net/publication/261323344

Paulston, C. B. (1974). Linguistic and communicative competence. *TESOL Quarterly*, *8*(4), 347-362. https://doi.org/10.2307/3585467

Pennycook, A. (1989). The concept of method interested knowledge, and the politics of language teaching. *TESOL Quarterly*, *23*(4), 589. https://doi.org/10.2307/3587534

Petersen, J. (2022). Using technology to assess interactive skills in a speaking test. *JALT Postconference Publication*, *2021*(1), 217. https://doi.org/10.37546/jaltpcp2021-25

Phan, S. (2008). Communicative language testing. *TESL Working Paper Series, 6*(2), 1-10.

Popham, W. J. (2008). *Transformative assessment*. Association for Supervision and Curriculum Development.

Priestley, M., Biesta, G. J. J., and Robinson, S. (2019). *Teacher agency: An ecological approach*. Bloomsbury Academic.

Rea-Dickins, P., and Germaine, K. (1992). *Evaluation*. Oxford University Press.

Rendle-Short, J. (2006). Alan D., and Catherine E. (eds.), The handbook of applied linguistics. *Language in Society*, *35*(5), 763–768. doi:10.1017/S0047404506310342

Rezalou, A., and Yağiz, O. (2023). The effects of communicative activities on EFL learners' speaking performances in CLT classroom. *International Journal of Languages' Education and Teaching*, *11*(1), 1–19. https://doi.org/10.29228/ijlet.65008

Richards, J. C. (2006). *Communicative language teaching today*. Cambridge University Press.

Richards, J. C., and Rodgers, T. (2001). *Approaches and methods in language teaching*. Cambridge University Press

Rivers, W. M. (1973). From linguistic competence to communicative competence. *TESOL Quarterly*, *7*(1), 25. https://doi.org/10.2307/3585507

Roever, C., and McNamara, T. (2006). Language testing: the social dimension. *International Journal of Applied Linguistics*, *16*(2), 242–258. https://doi.org/10.1111/j.1473-4192.2006.00117.x

Rouf, M. A., and Sultana, K. N. (2018). Learner role in CLT: Practices in the higher secondary classrooms. *Journal of NELTA*, *20*(1), 34–40. https://doi.org/10.3126/nelta.v20i1-2.19775

Ruiz, R. (1984). Orientations in language planning. *NABE Journal*, *8*(2), 15–34. https://doi.org/10.1080/08855072.1984.10668464

Saharani, P. S., Sahib, N., and Syam, U. (2023). Speaking anxiety in secondary school (A case study in junior high school). *The Academic: English Language Learning Journal*, *7*(2), 146–154. https://doi.org/10.52208/aellj.v7i2.301

Saldaña, J. (2021). *The Coding Manual for Qualitative Researchers* (4th ed.). SAGE.

Savignon, S. J. (1987). Communicative language teaching. *Theory Into Practice*, *26*(4), 235–242. https://doi.org/10.1080/00405848709543281

Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing*, *30*(3), 309–327. https://doi.org/10.1177/0265532213480128

Schaefer, M. B., Pennington, S. E., Divoll, K., and Tang, J. H. (2024). A systematic review of literature on student voice and agency in middle grade contexts. *Education Sciences, 14*(11), 1158. https://doi.org/10.3390/educsci1411115

Schreier, M. (2012). *Qualitative content analysis in practice*. SAGE Publications.

Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagné, and M. Scriven (Eds.), *Perspectives of Curriculum Evaluation*. Rand McNally.

Seidlhofer, B. (2005). English as a lingua franca. *ELT Journal*, *59*(4), 339–341. https://doi.org/10.1093/elt/cci064

Semiz, Ö., and Odabaş, K. (2016). Turkish EFL teachers' familiarity of and perceived needs for language testing and assessment literacy. *3rd International Linguistics and Language Studies Conference*, LILA'16

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information, 22*(2), 63–75. https://doi.org/10.3233/EFI-2004-22201

Shepard, L. A. (2000). The role of assessment in learning culture. *Educational Researcher*, *29*(7), 4-14. https://doi.org/10.3102/0013189X029007004

Shevelova-Harkusha, N. (2023). The use of communicative activities for developing speaking skills of students at english lessons. *ScienceRise: Pedagogical Education*, *2*(53), 20–25. https://doi.org/10.15587/2519-4984.2023.279581

Shohamy, E. (2017). Critical language testing. In E. Shohamy, I. G. Or, and S. May (Eds.), *Encyclopedia of language and education: Language testing* (3rd ed., pp. 441–454). Springer.

Sönmez, G., and Köksal, O. (2022). A critical overview of English language education policy in Turkey. *Shanlax International Journal of Education*, *11*(1), 1–9. https://doi.org/10.34293/education.v11i1.5721

Spillane, J. P. (2004). *Standards deviation: How schools misunderstand education policy*. Harvard University Press.

Spillane, J. P., Reiser, B. J., and Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research, 72*(3), 387–431. https://doi.org/10.3102/00346543072003387

Sreehari, P. (2012). Communicative language teaching: Possibilities and problems. *English Language Teaching*, *5*(12). https://doi.org/10.5539/elt.v5n12p87

Syarief, K. (2005). Communicative language teaching: Exploring theoretical foundations and practical challenges. *Jurnal Ilmu Pendidikan*, *12*(3), 1-14.

Şener, S. (2018). Teacher trainees' perceptions on English language speaking anxiety in foreign language classes. *Trakya University Journal of Education Faculty*. https://doi.org/10.24315/trkefd.303860

Tao, L., and Gao, X. (2017). Teacher agency and identity in curriculum reform: The case of Chinese secondary teachers. *Teaching and Teacher Education, 63*, 346–355. https://doi.org/10.1016/j.tate.2017.01.010

Taylor, L. (2005). Washback and impact. *ELT Journal*, *59*(2), 154–155. https://doi.org/10.1093/eltj/cci030

Tercan, G., and Dikilitaş, K. (2015). Efl students' speaking anxiety: A case from tertiary level students. *ELT Research Journal*, *4*(1), 16-27.

Tokçalar, F. (2025). Foreign language teaching and teacher training policies from the Ottoman period to the present. Anadolu University Faculty of Education Journal, 9(1), 91–100. https://doi.org/10.34056/aujef.1489386

Tomlinson, C. A. (2017). *How to differentiate instruction in academically diverse classrooms* (3rd ed.). ASCD.

Tsagari, D. (2009). The complexity of washback: An empirical study. *Frankfurt am Main: Peter Lang*.

Tsagari, D., and Banerjee, J. (2022). Language assessment literacy: Past, present and future. *Language Testing, 39*(2), 165–176. https://doi.org/10.1177/02655322211058077

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching*, *40*(3), 191–210. https://doi.org/10.1017/s0261444807004338

Veenstra, R. (Jay), and Weaver, C. (2022). Examining a continuum of FL speaking anxiety over time in an EFL classroom in Japan. *System*, *110*, 102889. https://doi.org/10.1016/j.system.2022.102889

Vogt, K., and Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, *11*(4), 374–402. https://doi.org/10.1080/15434303.2014.960046

Vorwerg, C. (2015). *Communicative Competence: Linguistic Aspects*, 294–301. https://doi.org/10.1016/b978-0-08-097086-8.53042-6

Watanabe, Y. (2004). Methodology in washback studies. In L. Cheng, Y. Watanabe, and A. Curtis (Eds.), *Washback in language in testing: Research contexts and methods* (1st ed., pp. 19-36). Routledge.

Weigle, S. C. (2002). *Assessing Writing*. Cambridge University Press.

Weir, C. J. (1993). *Understanding and Developing Language Tests*. Prentice Hall.

Wesche, M. B. (1983). Communicative testing in a second language*. *The Modern Language Journal*, *67*(1), 41–55. https://doi.org/10.1111/j.1540-4781.1983.tb02499.x

Wiggins, G. P. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. Jossey-Bass Publishers.

Winna, W., and Sabarun, S. (2023). The Language Assessment in Teaching-learning English. *DIAJAR: Jurnal Pendidikan Dan Pembelajaran*, *2*(4), 413–419. https://doi.org/10.54259/diajar.v2i4.1894

Zerey, Ö. G. (2008). *Impact of theater production on ELT students' foreign language speaking anxiety* [Master's Thesis, Mustafa Kemal University].

APPENDICES

APPENDIX 1
TEACHER INTERVIEW QUESTIONS FORM

GÖRÜŞME SORULARI

Tarih:

Cinsiyet:

Yaş:

Merhaba, ben Said Yasir KARAMAN. İngilizce öğretiminde MEB'in 2023 yılında getirdiği, konuşma ve dinleme becerilerinin sınav notunun %50'sini oluşturduğu yeni değerlendirme sistemi hakkındaki öğretmen ve öğrenci görüş ve düşüncelerini değerlendirmek adına bir araştırma çalışması yapıyorum. Bu araştırma kapsamında öğretmenlerin ve öğrencilerin yeni sistemin uygulanışı sırasındaki gözlemleri, ortaya çıkan muhtemel sorunlar ve bu sorunlar; yeni sistemin avantaj ve dezavantajları hakkındaki görüşlerini almak istiyorum. Araştırma sonucunda elde edilen bilgilerle yeni sistemin İngilizce dersinde öğretmen ve öğrenciler açısından deneyimlerini göz önünde bulundurarak dil değerlendirilmesinin geliştirilmesine katkıda bulunmayı amaçlamaktayım.

Görüşmeye başlamadan önce sizleri şu konularda bilgilendirmek isterim:

•Görüşmenin yaklaşık 10 dakika süreceğini tahmin ediyorum.

•Bu görüşme boyunca paylaştığınız bilgiler gizli tutulacaktır. Açıklanan hiçbir bilgi üzerinde isimleriniz belirtilmeyecek ve elde edilen bilgiler hiç kimseye iletilmeyecektir.

•Görüşme sırasında istediğiniz zaman soru sormakta ve görüşmeyi sonlandırmakta özgürsünüz.

•Başlamadan önce sormak istediğiniz başka bir soru var mı?

• İzin verirseniz görüşme sırasındaki konuşmaları kaydetmeyi istiyorum. Bu şekilde hem zamanı daha etkili kullanabiliriz hem de görüşmeden sonra cevaplarınızı daha detaylı analiz edebilirim. Sizce de uygun mudur? Herhangi bir rahatsızlık duyduğunuz taktirde görüşme kaydını ve görüşme notlarını sizinle paylaşabilirim.

Bu araştırmaya katılmayı kabul ettiğiniz ve zaman ayırdığınız için teşekkür ederim.

GENEL SORULAR

1. Hangi üniversiteden mezun oldunuz?

2. Kaç senedir İngilizce öğretmenliği yapıyorsunuz?

a. Hangi seviyelerdeki sınıflara eğitim veriyorsunuz?

3. Sizce İngilizce öğreniminde hangi beceri (okuma, yazma, dinleme, konuşma) daha önemlidir?

1

GÖRÜŞME SORULARI

1. Yeni değerlendirme sisteminin sizce için avantajları nelerdir ve İngilizce eğitimine ve öğrenimine ne tür olumlu etkiler ve yenilikler katmıştır?

2. Yeni sistemde, eğer karşılaştıysanız, hangi zorluklarla karşılaştınız?

3. Karşılaştığınız zorluklara karşı hangi çözümleri ürettiniz, üretilebilecek muhtemel çözümler ve önerileriniz nelerdir?

APPENDIX 2
STUDENT SURVEY QUESTIONS FORM

Öğrenci Görüş ve Deneyimleri

Tarih:

Cinsiyet:

Yaş:

Merhaba, ben Said Yasir KARAMAN. İngilizce öğretiminde MEB'in 2023 yılında getirdiği, konuşma ve dinleme becerilerinin sınav notunun %50'sini oluşturduğu yeni değerlendirme sistemi hakkındaki öğretmen ve öğrenci görüş ve düşüncelerini değerlendirmek adına bir araştırma çalışması yapıyorum. Bu araştırma kapsamında öğrencilerin yeni sistemin uygulanışı sırasındaki gözlemleri, ortaya çıkan sorunları, ve yeni sistemin avantaj ve dezavantajları hakkındaki görüşlerini almak istiyorum. Araştırma sonucunda elde edilen bilgilerle yeni sistemin İngilizce dersinde öğretmen ve öğrenciler açısından deneyimlerini göz önünde bulundurarak dil eğitimi ve değerlendirilmesinin geliştirilmesine katkıda bulunmayı amaçlamaktayım.

- Bu anket yaklaşık olarak 15 dakika sürecektir
- Bilgileriniz ve cevaplarınız gizli tutulacaktır. Açıklanan hiçbir bilgi üzerinde isimleriniz belirtilmeyecek, elde edilen bilgiler hiç kimseye iletilmeyecek ve hiç kimseyle paylaşılmayacaktır.

Araştırmaya katılmayı kabul ettiğiniz ve zaman ayırdığınız için teşekkür ederim.

Öğrenci Görüş ve Deneyimleri

1. Yeni değerlendirme sistemi sınıf içerisinde İngilizce öğrenme sürecinizi nasıl etkiledi? Eğer varsa, hangi yeniliklerle karşılaştınız örnek verebilir misiniz?

2. Yeni ölçme ve değerlendirme sistemi İngilizce konuşma ve dinleme becerilerinizi kullanmanıza nasıl yardımcı oldu ve bu becerilerinize ne tür bir katkıda bulundu?

3. Yeni ölçme ve değerlendirme sisteminde hangi zorluklarla karşılaştınız?

APPENDIX 3
PARENTAL CONSENT FORM

Sayın Veli;

Çocuğunuzun katılacağı bu çalışma, "Millî Eğitim Bakanlığının İngilizce Derslerine Yönelik Yeni Ölçme ve Değerlendirme Yönergesi uygulamalarına yönelik Öğretmen ve Öğrenci Deneyimlerinin İncelenmesi" adıyla yapılacak bir araştırma uygulamasıdır.

Araştırmanın hedefi Millî Eğitim Bakanlığı tarafından 11 Ekim 2023 tarihli yönetmelik ile getirilen yabancı dil değerlendirme değişikliğine geçiş dönemi ve bu yeni değerlendirme sisteminin İngilizce derslerinde uygulanış sürecindeki öğretmen ve öğrenci deneyimlerinin incelenmesidir. Araştırma uygulaması içersinde açık uçlu sorular bulunan anket şeklindedir.

Araştırma T.C. Millî Eğitim Bakanlığı'nın ve okul yönetiminin de izni ile gerçekleşmektedir. Araştırma uygulamasına katılım tamamıyla gönüllülük esasına dayalı olmaktadır. Çocuğunuz çalışmaya katılıp katılmamakta özgürdür. Araştırma çocuğunuz için herhangi bir istenmeyen etki ya da risk taşımamaktadır. Çocuğunuzun katılımı **tamamen sizin isteğinize bağlıdır**, reddedebilir ya da herhangi bir aşamasında ayrılabilirsiniz. Araştırmaya katılmamama veya araştırmadan ayrılma durumunda öğrencilerin akademik başarıları, okul ve öğretmenleriyle olan ilişkileri etkilemeyecektir.

Çalışmada öğrencilerden kimlik belirleyici hiçbir bilgi istenmemektedir. Cevaplar tamamıyla gizli tutulacak ve sadece araştırmacılar tarafından değerlendirilecektir.

Uygulamalar, genel olarak kişisel rahatsızlık verecek sorular ve durumlar içermemektedir. Ancak, katılım sırasında sorulardan ya da herhangi başka bir nedenden çocuğunuz kendisini rahatsız hissederse cevaplama işini yarıda bırakıp çıkmakta özgürdür. Bu durumda rahatsızlığın giderilmesi için gereken yardım sağlanacaktır. Çocuğunuz çalışmaya katıldıktan sonra istediği an vazgeçebilir. Böyle bir durumda veri toplama aracını uygulayan kişiye, çalışmayı tamamlamayacağını söylemesi yeterli olacaktır. Anket çalışmasına katılmamak ya da katıldıktan sonra vazgeçmek çocuğunuza hiçbir sorumluluk getirmeyecektir.

Onay vermeden önce sormak istediğiniz herhangi bir konu varsa sormaktan çekinmeyiniz. Çalışma bittikten sonra bizlere telefon veya e-posta ile ulaşarak soru sorabilir, sonuçlar hakkında bilgi isteyebilirsiniz. Saygılarımızla,

Velisi bulunduğum ..... sınıfı ..... numaralı öğrencisi .... 'in yukarıda açıklanan araştırmaya katılmasına izin veriyorum.
(Lütfen formu imzaladıktan sonra çocuğunuzla okula geri gönderiniz*).

...../...../.....

Öğrenci Adı-Soyadı ve İmza :

Veli Adı-Soyadı ve İmza :

Veli Telefon Numarası :

Araştırmacı : Said Yasir Karaman

İletişim bilgileri :

APPENDIX 4
VOLUNTARY CONSENT FORM

T.C.

ÇANAKKALE ONSEKİZ MART

ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

ETİK KURULU'NA

BİLGİLENDİRİLMİŞ GÖNÜLLÜ ONAM METNİ

Sizi Said Yasir KARAMAN tarafından yürütülen "Exploring English Teachers' and Students' Experiences on the Implementation of MoNE's New Assessment and Evaluation Regulation towards English Language Courses" başlıklı araştırmaya davet ediyoruz. Bu araştırmanın amacı Millî Eğitim Bakanlığı tarafından 11 Ekim 2023 tarihli yönetmelik ile getirilen yabancı dil değerlendirme değişikliğine geçiş dönemi ve bu yeni değerlendirme sisteminin İngilizce derslerinde uygulanış sürecindeki öğretmen ve öğrenci deneyimlerinin incelenmesidir. Araştırmada sizden tahminen 15 dakika ayırmanız istenmektedir. Araştırmaya sizin dışınızda tahminen 40 kişi katılacaktır. Bu çalışmaya katılmak tamamen **gönüllülük** esasına dayanmaktadır. Çalışmanın amacına ulaşması için sizden beklenen, bütün soruları eksiksiz, kimsenin baskısı veya telkini altında olmadan, size en uygun gelen cevapları içtenlikle verecek şekilde cevaplamanızdır. Bu formu okuyup onaylamanız, araştırmaya katılmayı kabul ettiğiniz anlamına gelecektir. Ancak, çalışmaya katılmama veya katıldıktan sonra herhangi bir anda çalışmayı bırakma hakkına da sahipsiniz. Bu çalışmadan elde edilecek bilgiler tamamen araştırma amacı ile kullanılacaktır.

Sıra No	Araştırmacının/Yürütücünün Adı Soyadı	Araştırmacının/Yürütücünün E- Posta Adresi
1	Said Yasir KARAMAN	
2	Dr. Öğr. Üyesi Sedat BECEREN	

☐ Araştırmaya katılmayı kabul ediyorum.

☐ Araştırmaya katılmayı kabul etmiyorum.

APPENDIX 5
ETHICS COMMITTEE APPROVAL

T.C.
ÇANAKKALE ONSEKİZ MART ÜNİVERSİTESİ REKTÖRLÜĞÜ
Lisansüstü Eğitim Enstitüsü Etik Kurulu
Bilimsel Araştırma ve Yayın Etiği

Sayı : E-84026528-050.99-2500009179 10.01.2025
Konu : Başvuru İncelenmesi

Sayın Said Yasir KARAMAN

Yürütücülüğünüzü yapmış olduğunuz 2024-YÖNP-5481 nolu projeniz ile ilgili Bilimsel Araştırma ve Yayın Etiği Komisyonu'nun almış olduğu 09.01.2025 tarih ve 01/65 sayılı kararı aşağıdadır.

Bilgilerinize rica ederim.

KARAR 65- Sorumlu yürütücülüğünü **Dr. Öğr. Üyesi Sedat BECEREN**'in yaptığı ve proje araştırmacısı **Said Yasir KARAMAN** tarafından gerçekleştirilen "Milli Eğitim Bakanlığının İngilizce Derslerine Yönelik Yeni Ölçme ve Değerlendirme Yönergesi Uygulamalarına Yönelik Öğretmen ve Öğrenci Deneyimlerinin İncelenmesi" başlıklı araştırmanın, Bilimsel Araştırma ve Yayın Etiği Komisyonu ilkelerine **uygun** olduğuna oy birliği ile karar verilmiştir.